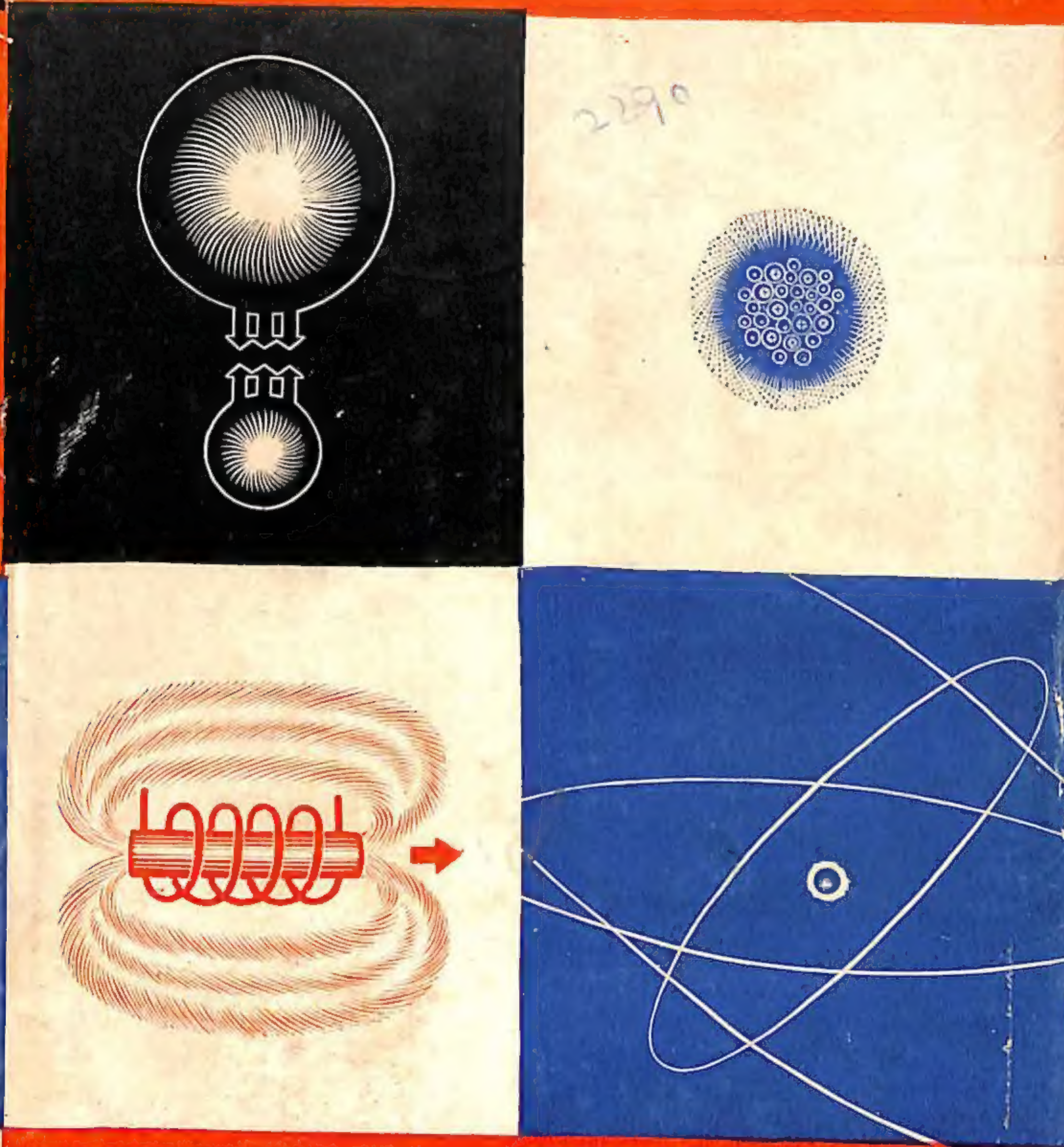
THE FORCES

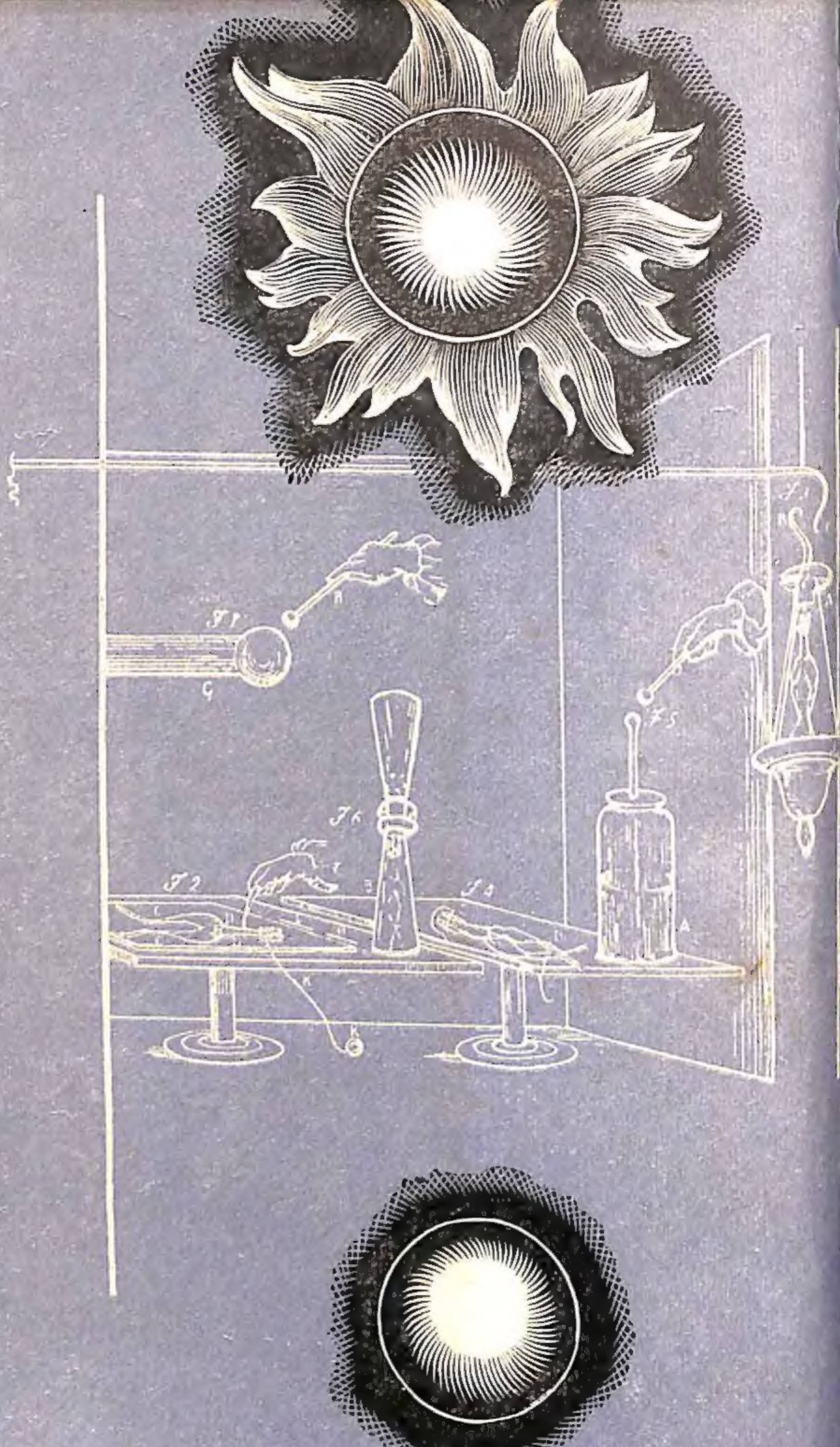


2290

OF NATURE

2290

MIR PUBLISHERS



V. GRIGORYEV
G. MYAKISHEV

THE FORCES OF NATURE

Translated from the Russian
by
George Yankovsky

MIR PUBLISHERS · MOSCOW 1967

UDC 530.1 (0.23) = 20

В. ГРИГОРЬЕВ, Г. МЯКИШЕВ

СИЛЫ В ПРИРОДЕ

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

Revised from the 1964 Russian Edition

На английском языке

CONTENTS

Chapter with No Number

ELECTROMAGNETIC FORCES IN ACTION

Chapter Four

NUCLEAR FORCES

Chapter Five

WEAK INTERACTIONS

Chapter Six

IN LIEU OF AN APPENDIX

C H A P T E R O N E

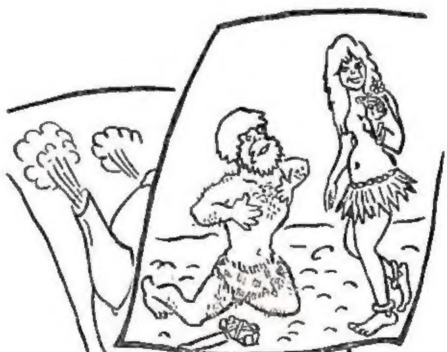
"Words, words, words."
Hamlet (Shakespeare)

IN LIEU OF AN INTRODUCTION

1

A Word About the Word "Force"

Let us first see what the dictionary has to say about the word "force". It states: "One of the separate causes to which classes of physical phenomena, as electricity, gravitation, and heat, were formerly ascribed." Yes, at one time the world was driven hard to explain all kinds of new



phenomena that were discovered. We find refractive forces, electrical forces, motive forces. The medieval scholastics explained temperature changes by means of a hot and a cold force.

It fell to physics to give "force" a rigorous definition. But even so the matter was not so simple as one might think, as we shall see as we go along.

2

The world as Newton saw it

The classical mechanics of Galileo and Newton became the cradle of the scientific meaning of the term force.

Generations of scientists from Newton's time to the present have been amazed by the magnificent integral conception of the world that was built up on the basis of Newton's works. "Mortals, congratulate yourselves that so great a man has lived for the honour of the human race." (Translation from the Latin of inscription on Newton's tombstone.)

According to Newton, the entire world consists of solid, ponderable, impenetrable, mobile particles. These primary particles are absolutely solid: they are immeasurably more solid than the bodies made from them; so solid that they are never worn out and never break into pieces. Such particles were believed, in the main, to differ quantitatively. The vast wealth and qualitative diversity of the universe is the result of different motions of these particles. In this world concept, motion is the most fundamental factor. The inner essence of the particles of matter is of secondary importance. The main thing is *how* these particles move.

Newton's laws of motion

The foundation of this unified picture of the world lies in the universal character of the laws of motion that Newton discovered and expressed in rigorous mathematical form. With remarkable exactitude these laws are obeyed by enormous celestial bodies and by the tiniest dust particles whipped about in the wind. And even the wind—which is the motion of invisible particles of air—obeys these same laws.

The central idea in Newton's laws is this: *Any change in the state of motion (that is, velocity) of bodies is due to their mutual interaction.*

Isn't this self-evident? No, not in the least. Newton, following Galileo, put the quietus to one of the greatest delusions of man concerning the laws of motion of bodies. Beginning with Aristotle, for nearly twenty centuries it was firmly believed that motion at a uniform speed could only be maintained through action from without, via some active cause, that any body would definitely come to a standstill if not so sustained.

This would seem to be corroborated by our everyday experience. For example, switch off the engine of an automobile and it will come to a stop on a perfectly horizontal road. Other conditions being equal, the velocity of a car increases with the power the engine develops. The same may be said of a boat, bicycle, steamship, and so on. That is why even in our day and age one comes across people who reason like Aristotle, though they do not realize it.

In reality we know that an isolated body, that is, one not interacting with any other body, always moves with a constant velocity. We often say that the body is moving by inertia. Only some action due to another body is capable of changing its velocity. The only reason why an effort is needed to maintain constant velocity in ordinary conditions is that there is always resistance offered by the ground, air or water, producing friction. If there were no friction, the speed of a car would not diminish when the engine is turned off.

This very thing was too much for the thick-headed and garrulous Colonel Kraus von Zillergut whose pincher

Schweik had stolen. "When the petrol came to an end," he said, "the car stopped. I saw it myself yesterday. And to think, gentlemen, people go on talking about inertia. It didn't go, it just stood there. No petrol. Isn't that funny, now?"

The most remarkable thing in Newton's laws of motion is their precise quantitative form. Not only can we speak about a certain interaction of bodies, we can measure the interaction. In mechanics, this quantitative measure of interaction of bodies is known as force.

What do muscular force and gravitation have in common?

Now a given body can be acted upon in a variety of ways. What, for instance, does the force of the earth's attraction to the sun have in common with the force that moves a rocket overcoming gravitation? Or compare these two forces with ordinary muscular strength. They are all quite different by nature. But perhaps they are in some way related physically. Newtonian mechanics says that they are. Here again we see that mechanics is nothing more than generalized everyday experience.

When a person tries to lift a heavy weight, he is actually comparing two quite different forces: muscular strength and the force with which the earth attracts the weight. But if you raise a heavy weight and hold it, you

can safely state that the muscular strength of your arms is equal in magnitude to the force of gravity.

This last assertion is essentially a definition of equality of forces in mechanics. Two forces, irrespective of their nature, are considered equal and opposite if their joint action on a body does not alter its velocity. This opens the way to a comparison of forces; and if one of them is arbitrarily chosen as a sort of standard, we can measure the forces.

Inertia

Please note that the most important thing in our definition of force is its association with motion. If a body is at rest, the forces acting upon it balance each other. The state of motion of a body changes only when the forces acting on that body are not in equilibrium, do not balance. The body is then accelerated, the magnitude of acceleration being directly proportional (according to Newton's laws of motion) to the *magnitude of the force*, but utterly *independent of the origin of the force*. The change in speed of a body is dependent not only on the force but on the body itself. The property of a body that determines the rate of change of velocity due to a force is in mechanics called mass (or inertial mass). According to Newton's second law, the acceleration (change of velocity with time) of a body is proportional to the force acting on it and is inversely proportional to the mass.

Thus, classical mechanics gives a rigorous definition of the term force and supplies a method for measuring forces. The action of a force is related to acceleration in a strictly quantitative fashion. Mechanics says Engels, is the only science in which "we know exactly what the word force signifies".

But even in mechanics the situation as regards forces can hardly be called satisfactory. We still don't know what physical processes give rise to a particular force. Apparently, Newton felt the same way. Here is what he had to say:

"I do not know what I may appear to the world, but, to myself, I seem to have been only like a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or a prettier shell than or-

dinary, whilst the great ocean of truth lay all undiscovered before me."

In mechanics, the difficulties associated with the nature of forces are ordinarily considered inessential simply because they are not spoken about. There is nothing wrong with this approach, for when we calculate trajectories of bodies it is sufficient to know the magnitude of the force. Now this can be found and we can determine when and how the forces operate without at all going into the nature of the forces and resorting solely to techniques of measuring them. For this reason mechanics "does not require that the definition of force should explain what a force is and whether it is the cause or the effect of the motion" (Henri Poincaré).

The fact that the nature of forces is not essential to mechanics is at once a defect and a merit. It is precisely for this reason that mechanics deals so successfully with the motion of molecules and stars.

This is remarkable. Yet the blotch remains. No wonder, then, that scientists felt the lack of a clear conception of force and made constant attempts to surmount such difficulties. This was done by passing from a somewhat formal introduction of forces to attempts at a more profound analysis of the nature of interactions. There were others who, like the famous German physicist Hertz, dispensed completely with the notion of "force" in mechanics.

Mechanics without forces and forces without mechanics

Hertz succeeded in constructing a mechanics without using the force concept at all. But the game was not worth the candle. In excluding force from mechanics it was necessary to introduce new hypotheses; however, this complicated the formulation of the basic propositions of mechanics. As a result the Hertzian scheme was not accepted.

The interesting thing was that this obscure understanding of the nature of forces which gave rise to attempts to rid science of forces altogether had just the opposite effect. The term force was extended from mechanics to other fields of science, losing on the way the rigour that

it had acquired within the framework of mechanics. F. Engels wrote in *Dialectics of Nature* that "if a cause of motion is termed a force, this does no damage to mechanics as such; but it becomes the custom to transfer this term also to physics, chemistry, and biology, and then confusion is inevitable." Quite some time has passed since these lines were written. Physicists have, in the main, gotten over such proclivities. However, remnants of that distant period that Engels wrote about still linger on in present-day terminology. Take, for instance, electromotive force (which in actuality is not a force but work); it has no relation to force in the ordinary mechanical sense of the word.

3

Robert Mayer and love of mankind

We have already spoken about how attempts to rid mechanics of forces failed. However, although force was retained in classical mechanics, the development of physics on the whole demonstrated that not every interaction by far could be completely described by mechanical forces.

At first it was hard to see any threat to our concept of force. Newtonian mechanics continued to develop. Other concepts were introduced: quantity of motion (momentum), energy, and so on. Gradually, energy (though the term was not originally used) began to acquire more and more significance. Like force, energy could describe the interaction of bodies quantitatively. In addition to interaction, it described the state of their motion. In mechanics, energy is determined both by the velocities of bodies and by the character of their interaction, which is of special importance to us. What is more,

it was found that all the basic propositions of Newtonian mechanics could be translated into the language of energy. Both descriptions of motion—in terms of force and in terms of energy—are equivalent (if we disregard forces that depend on velocity, for instance, friction). The work of a force is equal to the change in energy. Now the energy of a system of bodies may be regarded as the reserve of work that the system can perform. Generally speaking, the quantity of mechanical energy of an isolated system does not remain constant: friction reduces it.

The situation changed radically in the middle of the 19th century when the most fundamental law of modern science—the law of conservation of energy—was stated in precise form.

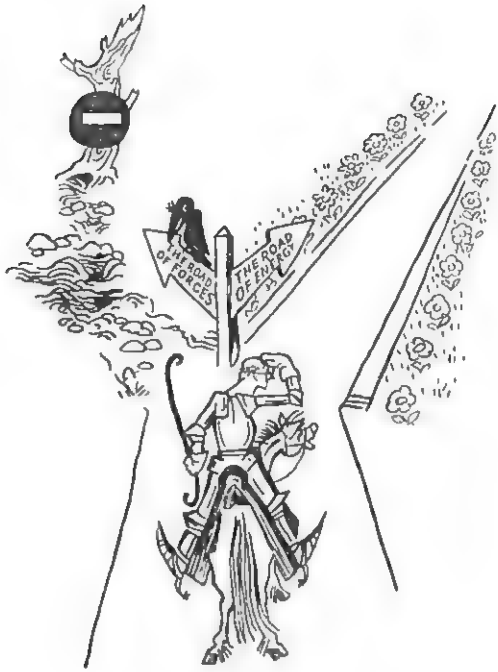
One of the founders of this law, Robert Mayer, gave the following appraisal: "Now, to have access to the science of motion, one does not need to scale the heights of mathematics first; on the contrary, nature itself reveals its beauty in full splendour, and even a person with meagre abilities can see a multitude of things that have hitherto been hidden from the greatest minds."

The law of conservation of energy states that the energy involved in mechanics does not disappear under any conditions. It can only be transformed from one form into another. When energy in the mechanical form disappears, the very same amount of energy of a different kind comes into existence. One instance is the heating of bodies.

Translating "egg"

Energy turned out to be a universal quantitative characteristic of motion and the interaction of all bodies—from astronomical objects to elementary particles. The law of conservation and transformation of energy is not confined to mechanical motion, and it is therefore no surprise that describing interactions in terms of energy is more general than with the aid of forces. Within the framework of Newtonian mechanics, however, neither one is predominant.

The transition of energy from one form to another may be likened to the translation of words from language to language. The translation changes the outward appearance



and the sound of the word, but the concept lying behind it remains. "Egg" is "das Ei" in German and "l'oeuf" in French, yet the object is the same.

In a way, forces are like idiomatic expressions—specific to the language in which they were born and meaningless in a word for word translation into another language.

It is interesting to note that originally, before the term "energy" came into established use, the word "force" was used to designate this new physical quantity. Helmholtz, whose name (along with Mayer and Joule) is associated with the discovery of the law of conservation of energy, entitled his work *On the Conservation of Force*. Mayer wrote: "If we once allow the word "force" to be used in two senses, it will later be a Herculean task to distinguish them in all particular cases." All his life Mayer insisted that the word "force" should be retained only for what we now call "energy". One can easily imagine what a confusion of tongues would have resulted.

Due to its great universality, the concept of energy could not but lead to the replacement of force descriptions by energy descriptions. The force concept has precise quantitative meaning only in mechanics, while that of energy embraces processes of any nature: we have thermal energy, electromagnetic energy, nuclear energy, and so forth.

When Newton's laws go on strike

The Newtonian description of motion is adapted to cases where relatively simple forces lead to motions that may prove rather complex. To illustrate, the extremely simple forces of universal gravitation lead to complicated planetary movements if one takes into account not only their relationship with the sun but with each other. Imagine that you pass from a description of the motions of a small number of bodies to a study of the motions of hundreds, thousands, millions and more particles. It might appear that Newton's mechanics could, in principle, give an accurate description of such systems, in other words, that it could at any time determine both the position and velocity of any particle. Actually, the mechanical approach here is meaningless. A precise statement of the problem (to determine the initial positions and velocities of the particles and also to specify the forces and interactions between them) is no easier than its solution. And true enough, the behaviour of large assemblies of particles exhibits new types of laws that cannot be reduced to mechanics. These are laws of statistical physics.

The sum of forces in a glass of water

In statistical physics, no attempts are made to follow the motions of the individual particles. What is considered is the average behaviour of large aggregates of particles. Now, since energy is conserved, we have the right to speak of the mean energy possessed by the parti-

cles of a system. But the mean force of interaction of particles with one another is not retained and is meaningless when applied to a large assembly of particles.

The forces of interaction between separate pairs of particles of the system are equal in magnitude and directed towards each other. For this reason, *the total sum of the forces operating within a system* is, generally speaking, zero. We can only speak of the mean force that the *system* (gas in a cylinder, for example) *exerts on some other body* (for instance, a piston that cuts off the gas).

Another thing to bear in mind is that any change in the state of a substance due to external forces is always associated with a change in its energy. But this change is not equal to the work of the forces, as in mechanics. A tea kettle can boil away without exhibiting any "forces"—no mechanical work is performed.

Concerning a great argument

Long before science had any inkling of molecular motions and statistical descriptions, mechanicians were already thinking about what at first glance appeared to be a rather insignificant problem: Is the interaction between bodies instantaneous or is a definite interval of time required?

In the chapters on gravitational and electromagnetic forces we shall discuss in detail the debate between the adherents of "action at a distance" (which was understood as instantaneous interaction through empty space without any intermediate agency) and those who supported the idea of "close-range action", the opposite of the former view. Arguments that brought about a radical change in the debate were advanced when the theory of electromagnetic phenomena made its appearance. These arguments were in favour of close-range action. Here, for the first time, it was clear beyond the shadow of a doubt that the "message" of any charge or current is not delivered instantaneously over all possible distances, but that the transmission of interaction takes time.

In other words, a signal, no matter what kind, is propagated with a very large velocity, but not an infinite velocity. And as soon as this was proved, the question arose: What is to become of Newtonian mechanics? The point is that in mechanics, all actions are reciprocal. The force exerted on a table by a book is equal in magnitude and opposite in direction to that exerted by the table on the book. According to Newton's third law, to every action there is an equal and opposite reaction.

Now if one of two interacting charges is moved suddenly, the other one will not "feel" it for a time. It will still be experiencing the action of the old force, whereas the former charge, as soon as it has moved, will immediately come under the action of the altered forces. The action is thus not equal to the reaction.

This is no small point, and as we shall see later on, it is not accidental either. The crux of the matter is that the intermediary, which participates in the interaction of the charge (an electromagnetic field) is not a mechanical system, which means that it cannot be described in terms of Newtonian mechanics. This fact is extremely important.

The situation in the theory of electromagnetic phenomena is by no means an exception. *The present predominant view of close-range action, that is to say, interaction by means of fields, deals a decisive blow at "forces" as an instrument for describing interactions.*

You can't put a spring in an atom

Despite the complications that have come with the introduction of fields, Newton's mechanics performs marvelously when describing the motions of charged particles in specified electromagnetic fields. (Naturally, electromagnetic fields do not obey the laws of mechanics that govern only the motion of the bodies themselves.) But even this partial application of mechanics is by far not always possible.

In the world of elementary particles, one is not able, using forces, to describe the interaction not only of

large assemblies of particles but even of individual particles.

In mechanics, force is uniquely related to the acceleration of a body at a definite instant of time and at a specific point in space. Now when dealing with the motions of elementary particles, not only acceleration but even the velocity at a point is absolutely meaningless. An elementary particle, say an electron, cannot be regarded as a simple sphere of very small size. Though an electron does definitely move in space with time, this motion cannot be visualized as a translation in space along a line or trajectory. What is more, it is impossible to measure the force directly in the microworld, with minute scales, for instance. We can't use a spring to measure the force of interaction of, say, an electron and a nucleus.

Classical mechanics, and with it the concept of force, is in general not applicable to elementary particles. Using forces, it is not possible to describe with exactitude the interaction of elementary particles in atoms and atomic nuclei. Only an energy description is possible. Energy is such a universal concept that the law of conservation of energy extends to elementary particles as well; true, in more complicated form.

Force is an interaction

Still and all, atomic physics makes use of forces. One frequently hears about nuclear forces operating in the atomic nucleus, electromagnetic forces of interaction between electrons, and so forth. In these instances, we have a new, and, we hope, a final meaning of this remarkable word. These are no longer the forces of mechanics. *The term force becomes synonymous with interaction.* This is no longer a strictly definite quantity that may be measured or put into an equation to describe actual processes. It is simply a qualitative definition of a type of interaction, an indication of its nature.

Thus, modern science uses the word force in two meanings: in the meaning of a mechanical force (here it is an exact quantitative measure of interaction) and—more

frequently—simply as the presence of interaction of a definite kind, an exact quantitative measure of which only energy can be. When speaking of nuclear forces we have in view the second meaning of this word. It is fundamentally impossible to include nuclear forces in the framework of Newtonian mechanics.

True, it might be possible to get along without the use of force in this new meaning. In one way, it is a step backwards. But probably tradition is so great in the use of this word that it will remain. Words both in everyday life and in science live their own lives, and neither reasoning "wisely" nor legislative fiat can change the situation.

4

How many forces are there in nature?

The title *The Forces of Nature* is mostly concerned with the second meaning of the word "force" as used in modern science. However, in many cases later on in this book we shall also have in view the narrower "mechanical meaning".

Above all, we shall discuss the nature of forces, something that mechanics refused to do altogether. There immediately arises a problem of prime importance: How many different types of force, or types of interaction, are there in the world?

Today, when one speaks of the unity of nature, he usually means unity in the structure of matter: all bodies are built up out of a few distinct elementary particles. This, however, is only one aspect of the unity of nature. No less essential is another aspect. Despite the remarkable diversity of actions, and reactions of bodies, which in the final analysis reduce to the interac-

tions of elementary particles, there are presently held to be only four distinct types of force. Only four types throughout the vast reaches of the universe, on our planet, in any piece of material, in living organisms, in atoms, atomic nuclei, and, finally, in the mutual transformations of elementary particles. They are the force of gravitation, electromagnetic forces, nuclear forces and weak interactions (this list does not include the so-called "nonforce interactions" expressed in quantum mechanics by the Pauli principle). Only the first two types may be regarded as forces in the sense of Newtonian mechanics.

What is to be understood by these types of force? And why is the number so small when we can name quite a few distinct forces such as gravitation, friction, elasticity, electric, magnetic, and nuclear forces, etc.?

To get the answers you will have to read the whole book. But it is well to bear in mind that when we say there are only four types of force, that of course does not mean that all known processes can be accounted for by their action. Simply, at the present time we cannot point to any phenomenon which would require the introduction of new forces other than those already mentioned.

The unity of natural forces is intimately bound up with the unity of the structure of matter. Not only is one inconceivable without the other, but it would probably be more appropriate to say that they both express different aspects of the unity of the world exhibited in all things of nature. To the relatively small number of kinds of elementary particles there corresponds a still smaller number of types of interaction between them.

What this book will contain

Now about the most important thing. What are these types of interaction and how were they discovered? How is it possible to explain the infinite diversity of interactions of bodies using a few general laws? What is the sphere of action of the various forces in nature and what role do they play in the different processes? Then there

is the question of the interrelationship of forces, the harmony of natural forces which ensures a relative stability and constant development and regeneration of the universe, where all forces are equally needed.

We shall begin our study of natural forces where physics began—with the forces of universal gravitation. Gravitational forces stand at the head of an astounding chain of discoveries that led to the establishment of the unity of the forces of nature.

CHAPTER TWO

The stars streamed in arcs—
before us!
And the dawn shone brilliantly—
before us!

Rubaiyat (Omar Khayyam)

GRAVITATIONAL FORCES

1 *From Anaxagoras to Newton*

2 *The Law of Universal Gravitation*

3 *Gravitation in Action*

4 *Geometry and Gravitation*

1

The infinite fall

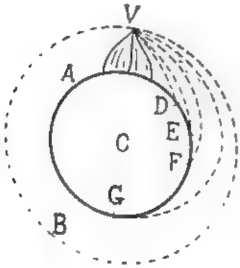
Figure 213 in Newton's great work *The Mathematical Principles of Natural Philosophy* is remarkable for the simplicity with which it illustrates the profound relationship between terrestrial and celestial mechanics. The caption to this figure states that a stone which has been thrown will deviate, under the action of gravity, from a rectilinear path, and will fall to earth describing a curved trajectory. If it is thrown with a greater velocity, it will fall farther away. Continuing this reasoning, Newton arrived at the conclusion that without air resistance and with sufficient velocity the trajectory would be such that the stone would never again come to earth and would revolve about the earth like a planet describing its orbit in outer space. This is all very topical when one recalls the numerous launchings of artificial earth satellites during recent years.

To summarize, then, the motions of planets round the sun and of our moon around the earth all represent a falling, which simply continues on and on without end (at any rate if one disregards the transformation of energy into "nonmechanical" forces). This "falling", whether in the form of a stone falling to earth or of planets moving round in their respective orbits, is due to the force of gravitation.

Long before Newton, there were conjectures about the unity of causes governing planetary movements and the falling of objects on earth. Probably the first to express this idea clearly was the Greek philosopher Anaxagoras from Asia Minor who lived in Athens about two thousand years ago. He said that if the moon did not move it would fall to earth like a stone hurled from a sling. Not so badly stated, especially if one recalls that the statement was made over twenty centuries before the time of Newton.

However, Anaxagoras' brilliant conjecture apparently had no practical effect on the development of science. It was destined to be forgotten by future generations and misunderstood by his contemporaries. Ancient and medieval thinkers who dealt with planetary motions were very much removed from a correct (or, in fact, any) interpretation of the causes of such movements. Even the great Kepler, who through titanic effort succeeded in formulating exact mathematical laws of planetary motion, believed that the underlying cause of this motion was the rotation of the sun.

According to Kepler, the sun, in rotating, pulls the planets round by constant impulses. True, it was not clear why the periods of revolution of the planets about the sun differed from the rotational period of the sun itself on



its axis. About this Kepler wrote: "...if the planets did not possess a natural resistance, it would be impossible to indicate the reasons why they do not follow precisely the sun's rotation. And though in reality all planets move in the same direction as the sun rotates, their velocity of motion is not the same. The point is that they combine in certain proportions the reluctance of their own mass with the velocity of their motion."

Kepler did not realize that the coincidence of the directions of planetary motion about the sun with the direction of solar rotation has nothing to do with the *laws of planetary motion* but is connected with the *origin* of our solar system. An artificial planet may even be launched in a direction opposite to solar rotation.

Much closer to the discovery of the law of attraction of bodies than Kepler was Robert Hooke, who in *An Attempt to Prove the Motion of the Earth from Observations* (1674) wrote:

"[At a future date] I shall explain a System of the World differing in many particulars from any yet known, [and] answering in all things to the common rules of mechanical motions. This depends upon three suppositions: first, that all celestial bodies whatsoever have an attraction or gravitating power towards their own centres, whereby they attract not only their own parts, and keep them from flying from them, as we may observe the earth to do, but that they do also attract all the other celestial bodies that are within the sphere of their activity;... The second supposition is this: that all bodies whatsoever that are put into a direct and simple motion, will so continue to move forward in a straight line, till they are by some other effectual powers deflected and bent into a motion, describing a circle, ellipse, or some other more compounded curve line. The third supposition is: that these attractive powers are so much the more powerful in operating, by how much the nearer the body wrought upon is to their own centres. Now what these several degrees are I have not yet experimentally verified; but it is a notion, which if fully prosecuted as it ought to be will mightily assist the astronomer to reduce all the celestial motions to a certain rule, which I doubt will never be done true without it."

The astounding thing is that Hooke himself did not undertake to develop these ideas because he was busy with other work.

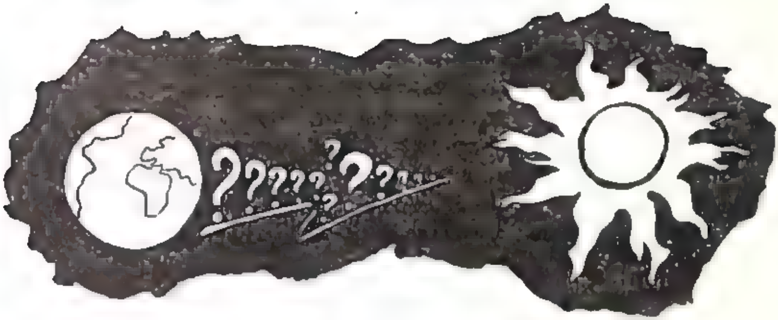
Newtonian mechanics and gravitation

The history of Newton's discovery of the law of universal gravitation is so well known that we hardly need to go into the details about how, when still a student, Newton first grasped the idea that the nature of the forces which make a stone fall and which govern the motion of celestial objects is the same, that the first calculations did not yield correct results because the available data on the distance between the earth and the moon were not exact, and that 16 years later new and improved measurements came to light. The theory was published only after new calculations were completed dealing with the lunar motions and those of all the known planets of the solar system, of comets, and the tides.

The discovery of the law of universal gravitation is rightly regarded as one of the greatest triumphs of science. This triumph is associated with the name of Newton. But why didn't Galileo, who had discovered the laws of free fall (and who incidentally spent much more time studying astronomy than did Newton), or Robert Hooke or some other of the outstanding predecessors or contemporaries of Newton make this discovery?

It was not fortuitous, nor did it have anything to do with "falling apples", nor even the "degree of genius", though this circumstance was undoubtedly very essential. The principal, decisive factor was that Newton had at his disposal the laws he discovered describing all possible motion. It was precisely these laws, which we now term Newtonian mechanics, that enabled Newton to grasp fully that forces underlie all phenomena and motions. Newton was the first to perceive with absolute clarity *the precise thing that had to be sought* in order to explain planetary motions. *The search was for forces, and only forces.*

Kepler established the exact trajectories of the planets of the solar system. He determined how the positions



of the planets in space vary with time. Given a definite trajectory, the equation of motion permits one straightway to determine the force that causes this motion. That was the problem which Newton solved.

What are these forces? What role do they play and what is their place in nature? And, finally, what of their physical origin?

The questions are many, and even today we do not have all the answers. That is for the physics of tomorrow. But much, and above all the law of universal gravitation itself which was so neatly formulated by Newton, has long since become the bedrock of science.

2

Acting on all things, it knows no barriers

One of the most remarkable properties of the forces of universal gravitation is reflected in the name that Newton gave them: *universal*. These forces are the most universal of all the forces of nature. Everything that has mass—and mass is inherent in all matter—must experience gravitational action. There is no exception even for

light. If we visualize gravitational forces in the form of strings stretching from one body to another, then all space could be imagined as permeated with numberless strings. And there is no way of "cutting" one of them and escaping the action of gravitational forces. Gravitation has no barriers. An electric field, by contrast, can always be cut off (say, by a screen of conducting material). A magnetic field is not able to penetrate into a superconductor. But gravitational effects are freely transmitted through all bodies. Shields made out of special materials impermeable to a gravitational force (like the "cavorite" of Wells' *First Men in the Moon*) can exist only in the imagination of science-fiction writers.

Quite recently, the French astronomer Allain made some measurements during a solar eclipse and was supposed to have found a "gravitational shade" and that the force with which the earth is attracted by the sun diminishes when the moon comes between them. Actually, however, the scientist simply forgot the temperature change of the instruments, which is inevitable during an eclipse. It was this insignificant, at first glance, effect that misled Allain.

How great are gravitational forces?

So gravitational forces exist everywhere and are all-penetrating. Then why don't we feel the attraction of most bodies? Why, for instance, do we feel the attraction of the earth at all times, while even the largest mountains attract only mountain climbers and eagles? Take Mt. Everest and calculate the portion of the earth's attraction under the most favourable location that it represents—it comes out to only thousandths of one per cent. The force of attraction of two medium-weight people at a distance of one metre does not exceed three hundredths of a milligram. That is how weak gravitational forces are. "Weak?" one perforce exclaims recalling how the earth is linked to the sun and the moon to the earth, and over enormous distances. Time and again these doubts have arisen. The well-known science popularizer Ya. Perelman mentions the appearance, at the end of

last century (not so long ago, as you see), of a book by Carpenter, *Modern Science*, where it was stated that the extreme weakness of gravitational forces claimed by physicists undermines the authority of physics.

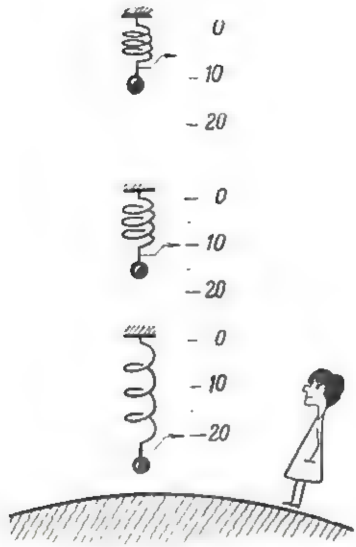
The fact that gravitational forces are, generally speaking, much weaker than electric forces brings about a peculiar division of the "spheres of influence" of these forces. For instance, the gravitational attraction of electrons (in atoms) to the nucleus is weaker than electrical attraction by a factor of 1,000,000,000,000,000,000, 000,000,000,000,000,000,000; it is easy to see that processes taking place inside the atom are determined almost solely by electrical forces (for the time being we disregard intra-nuclear processes). Gravitational forces become perceptible and at times fantastically powerful when we deal with such enormous masses as those of cosmic bodies: planets, stars and the like.

Thus, the earth and moon attract with a force of roughly 20,000,000,000,000,000 tons. Even very distant

stars whose light takes years to reach the earth "greet" us gravitationally to the extent of hundreds of millions of tons.

Their range of action is equal to infinity

We have taken it for granted that the mutual attraction of two bodies falls off as the distance between them increases. It is so obvious that no one seems to have any doubts on this score. Let us imagine the following experiment. We shall measure the force with which the earth attracts some body, say a twenty-kilogram weight. For the first experiment, let the weight be placed at a great distance from the earth. In these conditions the force of attraction (which, incidentally, may be measured with an ordinary spring balance) is practically zero. As the weight is brought closer to the earth, the mutual attraction will increase, and, finally, when it comes to rest on the surface, the balance will show "20 kg of force", because what we call weight (disregarding the earth's rotation) is nothing other than the force with which the earth attracts bodies. Now let us continue the experiment and lower our weight into a well. This will obviously diminish the force acting on the weight, for if we were to continue lowering it right down to the centre of the earth, the attraction from all sides would be in equilibrium and the pointer of the spring balance would stand at zero. What this shows, then, is that we cannot simply say that gravitational forces diminish with increasing distance. The statement has to be qualified to indicate that these distances must be much greater than the dimensions of the bodies. Only then will Newton's law hold: *the forces of universal gravitation diminish as the square of the distance between the attracting bodies.* Let us try to get a clearer picture of just what this means. Arithmetically speaking, this means that if the distance increases three times, the force will fall off by a factor of 3^2, or 9, and so forth. However, it is not yet clear whether this signifies a rapid or not so rapid variation of force with distance. Does the law mean that interac-



tion is perceptible only between "immediate neighbours" or is it appreciable at sufficiently large distances?

A convenient way to answer this question is to compare the above law with the law of diminishing light intensity as the light source recedes. In the case of both gravitation and light, the law is the same—they fall off in inverse ratio to the square of the distance. But we see stars at distances over which light—the fastest of all things—takes millions (!) of years to travel. Now if light from these stars reaches us, this should mean (since the law operates in the same way) that attraction will be felt, even though very feebly. Hence, the force of universal gravitation diminishes constantly with distance, but extends to practically unlimited distances. In the parlance of physicists, their range is equal to infinity. *Gravitational forces are long-range forces.* That is what they are called in physics. As we shall see later on, not all forces by far are of this nature. This long-range quality is what enables gravitation to link all the bodies of the universe.

The relative slowness with which these forces fall off with distance is felt all the time here on earth. When bodies change their altitude above the earth, they do not exhibit any change in weight (actually there *is* a change,

but it is practically negligible) precisely because gravitational forces change very slightly for small variations in distance (in our case, the distance to the centre of the earth).

Incidentally, it is for this reason that the law of variation of gravitational forces with distance was first discovered "in the sky". The facts came from astronomy. Even so, we can notice a change in the force of gravity right here on earth. For instance, a pendulum clock with a period of oscillation of one second will slow down about three seconds a day if raised from the basement to the topmost story of Moscow University (200 metres), and only because of the change in the force of gravity.

Altitudes at which artificial earth satellites move are comparable to the earth's radius, so that trajectory calculations have to take into account changes in the force of the earth's gravitation with distance.

An unusual property of gravitational forces

For centuries, medieval science accepted as an inviolable dogma Aristotle's contention that a heavier body falls faster. Everyday experience would seem to confirm this, since a feather obviously falls more slowly than a stone. However, it took Galileo to demonstrate that the crux of the matter lies in the resistance of the air, which radically alters the picture, distorting the law of falling bodies when the latter are acted upon by gravity alone. A convincing experiment may be performed with what is known as Newton's tube. It will permit us to evaluate properly the role of air resistance. Imagine an ordinary glass tube containing several things: shot, pieces of cork, feathers, fluff. When the tube is turned upside down, the heavy shot falls fastest, followed by the pieces of cork, with the fluff leisurely coming down last. Now pump the air out and see how these objects fall. Fluff, shot and cork now all fall at the same speed. Which means that the different speeds in the first experiment were due to the resistance of the air—greatest in the case of the fluff. Thus, we can say that if there were no air resistance and all bodies were subject solely to the

force of universal gravitation (the earth's gravity in our case), then all bodies would fall exactly with the very same rate of acceleration.

But "nothing is new under the moon". Two thousand years ago, Titus Lucretius Carus in his famous poem *De Rerum Natura* (*Of the Nature of Things*) wrote: "For all things that fall through the water and thin air, these things must needs quicken their fall in proportion to their weights, just because the body of water and the thin nature of air cannot check each thing equally, but give place more quickly when overcome by heavier bodies. But, on the other hand, the empty void cannot on any side, at any time, support anything, but rather, as its own nature desires, it continues to give place; wherefore all things must needs be borne on through the calm void, moving at equal rate with unequal weights." These remarkable words were of course only a beautiful conjecture. To transform this conjecture into a reliable law required numerous experiments: from the famous experiments of Galileo who studied bodies falling from the leaning tower of Pisa—balls of the same size but made of different materials (marble, wood, lead, etc.)—to intricate modern methods for measuring the effects of gravitation on light. This vast range of experimental findings is a convincing demonstration that gravitational forces impart the same acceleration to all bodies. As an illustration, the acceleration of free fall due to the earth's gravity is the same for all bodies and is not dependent either on the composition or structure or mass of the bodies involved.

This apparently simple law expresses perhaps the most remarkable peculiarity of gravitational forces. There are no other forces that accelerate all bodies to the same extent, irrespective of their mass. A football player kicks the ball. The lighter the ball, the greater velocity it acquires (for a blow of the same force and duration). What would you say about a kick that would accelerate in the same way a ball, a heavy weight and, say, an elephant? Incredible! Yet that is exactly how gravitational forces operate, the only difference being that the gravitational "kick" is continuous, never ceasing.

There is a profound physical significance underlying this marvelous peculiarity of gravitational forces, and

we shall have more to say on this question when discussing the nature of universal gravitation and the general theory of relativity. Let us now try to recall what lies at the heart of motion in mechanics. When we gave the definition of mechanical force, we had to rely on Newton's laws of mechanics, according to which the acceleration imparted to a body is directly proportional to the force acting on it and inversely proportional to the mass of the body. This leads to a simple and marvellous conclusion: for acceleration to be independent of the mass it is necessary that the force be proportional to the mass. By way of illustration, we take two bodies, a ping-pong ball and a lead sphere of the same size. The mass of the latter is 300 times that of the former. Which means that in order to accelerate the lead sphere in the same way as the ping-pong ball, we must act on it with a force 300 times as great. But the earth's gravity makes both objects fall with the same acceleration. Thus, this attraction has been "adjusted" in accord with the masses of the bodies: the greater the mass of the lead sphere as compared with that of the ball, so much stronger is the attraction of the earth.

Now, at last, we can compress into one succinct statement this remarkable property of gravitational forces: *the gravitational force is proportional to the masses of the bodies involved*. It is important to stress that we are

dealing with the mass which in Newton's laws appears as the measure of inertia. It is even called "inert mass".

The words "gravitational force is proportional to the mass" contain an amazingly profound meaning. Whether the bodies are large or small, cold or hot, whatever their chemical composition or structure, they all experience the same gravitational action (provided their masses are equal).

Perhaps this law is indeed simple, for Galileo considered it to be almost self-evident. Here is how he reasoned. Let two bodies of different weight be falling. On Aristotle's view, the heavier body will fall faster even in a vacuum. Now let us connect the two bodies. Then, on the one hand, the bodies should fall still faster because the total weight has increased. Yet, on the other hand, adding to a heavy body a portion that falls more slowly should slow it down. This is a contradiction that can be reconciled only when we make the assumption that all bodies acted upon solely by the earth's gravitational pull fall with the same acceleration.

Everything seems logically consistent, yet let us go over the reasoning once again. It is based on a frequently employed method called reductio ad absurdum: assuming that heavy bodies fall faster than light ones, we arrive at a contradiction. Note that from the very start it was assumed that the acceleration of free fall is determined by

the *weight* and only the *weight* (strictly speaking, it is not the weight but the mass).

Yet, a priori (that is, prior to experiment) this is not obvious at all. But suppose the acceleration were determined by the volumes of the bodies? Or by the temperature? Or—by a flight of the imagination—by the colour or smell? In short, let us imagine that there exists a gravitational charge similar to an electric charge and that, like the latter, it is in no way associated directly with the mass. This comparison with an electric charge is very useful. Take two particles of dust between the charged plates of a capacitor. Let the particles have equal charges but with mass in the ratio of 1 to 2. Then the accelerations should differ by a factor of two. The forces determined by the charges are equal, but for equal forces, a body of double mass will receive only one-half the acceleration. If the particles are combined, the acceleration will obviously have some new intermediate value. No purely speculative approach without an experimental study of electrical forces can yield anything. The picture would be

exactly the same if the "gravitational charge" were not connected with the mass. Only experiment can tell us whether there is such a connection or not. And so now it is clear that the experiments which proved the identity of acceleration of all bodies due to gravitation demonstrated that essentially the "gravitational charge" (a gravitational or heavy mass) is equal to the inert mass.

Experiment and only experiment can serve as the basis for physical laws and also as a criterion of their exactitude.

Experiment it is—from the small laboratory of the scientist to the immensity of the cosmic scale—that supports the law of universal gravitation, which (to bring into a focus everything that has been said) states:

The force of mutual attraction of any two bodies, whose sizes are much smaller than the distances between them, is proportional to the product of the masses of these bodies and is inversely proportional to the square of the distance between the bodies.

The proportionality factor here is called the gravitational constant. If we measure the distance in centimetres, the time in seconds, and the mass in grams, the gravitational constant will numerically be equal to 6.68×10^{-8} with dimensions $cm^3/g\ sec^2$.

3

The solar system

The words "celestial mechanics" have a sort of old-fashion ring to them. But celestial objects did not come into science by accident, they were the first ones on which the law of universal gravitation was tested. It was the captivating harmony of a unitary mathematical law governing the motions of the planets in their eternal whirling round the sun that first so forcefully attracted

physicists, astronomers and, generally, all natural scientists to the Newtonian theory.

One of the greatest triumphs of natural science is linked up with studies of planetary motions: the prediction of a new planet—Neptune—by the Frenchman Leverrier and the Englishman Adams. Slight deviations in the orbital motion of Uranus from values calculated on Newton's theory were interpreted as a perturbation due to some unknown planet. The orbit of this hypothetical planet was computed and when astronomers directed their telescopes at the indicated part of the sky, the new planet was straightway discovered.

Right up to the present day, universal gravitation remains in our minds as the basic motive force of cosmic bodies. It may seem strange that gravitational forces, which are so negligibly small in the interactions of bodies surrounding us, should play such a decisive role on the cosmic scale.

An incisive explanation to bring this fact home is given by Ya. Perelman. The masses of celestial bodies are truly enormous. But so also are the distances between such bodies. Now the force of gravitation is, as we recall, proportional to the product of the masses and inversely proportional to the square of the distance between them. The mass of a body is proportional to its volume, and hence, to the cube of its linear dimensions. For this reason, if the dimensions of bodies and their distance apart are increased n times, the force of gravitation will increase $\frac{n^3 \cdot n^3}{n^2} = n^4$ times! Which means that an increase in the dimensions of a portion of the universe by a factor of two would yield an increase in the gravitational forces by a factor of 16! That is why the attraction of cosmic bodies at large distances is incomparably greater than that of small bodies at close distances.

Your desk calendar indicates the phases of the moon, the time and type of solar and lunar eclipses. Astronomers publish extremely accurate tables indicating, for many years into the future, the exact positions of planets in the sky and the exact time of arrival in those positions. Very exact predictions are made of the arrival of many comets, or "long-haired" luminaries as they were called in the old days.

The origin of the planets

We seem to know everything about the motions of the planets of our solar system. But this is not so, and by a long shot. This system originated in the distant past, and has been developing and changing ever since, even now. It has a definite age. The stability and invariability of our planetary system over the centuries is only apparent, since people have observed it for a relatively short space of time as compared to its age. Observe a person for one tenth of a second, and he will appear to be a very "stable" system too.

What can we say about the past of our solar system? This question has long excited the curiosity of thinkers. Numerous hypotheses and conjectures have been advanced. Some are naive, some poetical, and others simply wild flights of the imagination.

We shall not touch on the numberless guesses of pre-Newtonian times. For the most part they were hardly scientific in the modern sense of this word.

First of all, where did the building materials for the solar system come from? More precisely stated, we want to know about the building materials of the planets, for scientists are on the whole agreed that the sun is older than its family of planets.*

This problem has been under discussion for a very long time. Each one of two principal hypotheses has its ardent adherents, and the balance moves up and down alternately, so that it is hard to draw any definitive conclusions.

The well-known Kant-Laplace hypothesis, which once appeared to be "almost obvious", states that the matter which went to form the planets was in the form of enormous incandescent "splashes" swung off the surface of the sun in rotation.

A modern version of this idea is that the building material of the planets appeared together with the sun, being a part that detached itself from the sun during the formative stages of the latter when it emerged from an

---

* The problem of the origin of stars, one of which is our sun, is of great interest in itself. But it is far from any solution and we shall not touch on it at all. Stellar processes will be mentioned in the chapters on nuclear forces and weak interactions.

interstellar condensation of gas and dust. Thus, Academician V. Fesenkov writes: "While not yet having formed into a star, that is, while continuing to contract intensively, the sun should have left a considerable quantity of matter approximately in the equatorial plane, which, due to the excessive speed of rotation, could not concentrate into a single body."

Another, and opposite, one is the "capture" hypothesis, which assumes that the building material of the planets came from interstellar space, the function of the sun being only to sweep up the necessary material and hold on to it. This notion was first clearly formulated by Academician Otto Shmidt of the Soviet Union, who was a noted mathematician, polar explorer, geographer, astronomer and geophysicist—other specialities could be added—and a remarkable man with a versatile range of knowledge.

Interstellar space contains enormous gaseous-dust clouds consisting of individual atoms, molecules and minute dust particles. They are familiar to astronomers and have been termed "coal bags". They prevent us from seeing many stars, including the centre of our own Galaxy. Matter in these clouds is in a very rarefied state, much more rarefied than the uppermost layers of the earth's atmosphere. But since they cover vast expanses, their overall mass is truly cosmic, really fantastic. These clouds are in motion. Therefore, it is entirely within the realm of possibility that a star could encounter such a cloud and sweep up a good deal of building material. That, say the capture-theorists, is how our solar system was born.

"But", their opponents object, "the probability of such capture is very small, since the sun's velocity relative to the clouds is very large (tens of kilometres a second). It would be necessary for the attraction of yet another star in the immediate vicinity to slow down their relative velocity. But as we know there is almost a negligible probability of such a stellar approach."

Let us not go further into the pros and cons. All the more so since there is yet no way of making a final choice between these hypotheses. Meanwhile, it may be noted that despite the essential difference between the theories of capture and joint formation (or "outflow"),

they have much in common. And that is what we shall talk about now.

No matter what the origin of the raw materials for our planets, they had to pass through numerous transformations before arriving at the present state of the solar system.

What picture should we start with? Most likely, a gigantic gaseous-dust cloud in rotation about the sun preceded the appearance of planets. Both hypotheses are in agreement with this assumption.

So it's a gigantic cloud

From then on the future of the cloud was determined in the main by three circumstances: the gravitational interactions of particles with the sun and among themselves, particle collisions, and, finally, the action of solar radiation. Rotation is also a factor naturally. The theory of the origin of the solar system developed by O. Shmidt, his pupils and coworkers (and this is a real theory and not simply a hypothesis) follows the evolution of the circumsolar cloud right up to the time of the formation of the planets. Putting into words the involved formulas that scientists manipulate, we get the following picture.

Through millions of years the cloud takes on multifarious shapes. Very slowly it begins to flatten out into a rotating disk. The particles come closer together and the forces of attraction increase. This produces numerous condensations and is the beginning of what is termed the gravitational condensation of matter. It somewhat resembles the formations of droplets of mist from water vapour. The forces here, however, are not molecular but gravitational.

Millenia pass before particle collisions in the dust condensation lead to the formation of solid bodies tens and hundreds of kilometres in size. Great numbers of these relatively large asteroidal bodies fill the gap between Mars and Jupiter.

For thousands of millions of years, these bodies move round the sun colliding, breaking up or sweeping up smaller bodies and fragments of larger ones.

Those that escaped fragmentation began to build up faster sweeping up debris scattered about in space. As

they became larger, they attracted with greater force and captured more particles. Truly cosmic intervals of time were needed for this condensation process to generate planets out of numberless chunks of material. About seven thousand million years have passed from the start.

Only in one place: near the most massive planet Jupiter, was the perturbing action so great as to prevent these chunks from coalescing. To this day we have a belt of asteroids here.

The Shmidt theory has even proved capable of accounting quantitatively for many regularities in the structure of the solar system. First, the law of planetary distances, which states that the square roots of orbital radii increase roughly as an arithmetic progression. The theory explains why the orbits of the planets are almost circular—very slightly elongated ellipses. Also explained is the fact that the orbital planes are but slightly inclined to each other and to the equatorial plane of the sun. The theory states that all planets should move round the sun in the same direction, which is absolutely true both for the major and minor planets. The theory likewise accounts for the fact that all the planets (with the exception of Uranus) rotate on their axes in the same direction. Finally, it very accurately predicted the distribution of masses and densities of objects of the solar system.

As the planets grow, gravitational forces should compress them steadily generating colossal pressures. At this point, the planets begin to heat up. However, the present high temperatures in the deep interior of the earth are due not only to formative processes. Heat released in the disintegration of radioactive elements (uranium, thorium, radium, etc.) gradually transformed the earth into a gigantic crucible which under enormous pressures wrought new materials that were squeezed to the surface where they cooled into a mantle that now covers the earth.

But that is not all. At the beginning of evolution of the solar system, dust particles collecting into a flat disk shut out the sun's rays to the outlying parts, where the cold of deep space reached down to 270°C below zero. At the same time, those portions of the disk that were closer to the sun were greatly heated by solar radiation.

As a result, the dominant particles close to the sun were of a refractory nature, while the gases (mainly hydrogen) became frozen in the cold outlying portions of the disk. Enormous masses of frozen gases became the building material of the more distant planets such as Jupiter and Saturn. These planets should be found to consist mainly of hydrogen, which becomes solid at the low temperatures out there.

On the contrary, the planets closer to the sun, like the earth, should be made up of refractory materials. One might be inclined to think that Uranus, Neptune and Pluto contain more hydrogen still than Saturn and Jupiter, but this is not so. The formation process of blobs of matter inside the dust cloud on the outskirts of the solar system was extremely slow because of the small density of the matter. The appearance of numerous asteroidal bodies near the sun increased the transmitting power of the cloud at a time when this process in the distant zone had not yet developed. Under such conditions, the sun's rays caused all the hydrogen to evaporate from dust particles and the surfaces of large pieces. This reduced the amount of hydrogen in the distant planets.

That is the theory. What about actuality? Do we find confirmation of these conclusions? We can't yet sample the soil of Jupiter, but knowing the mass and dimensions of a planet, it is possible to figure out the chemical elements it contains. Calculations show that Jupiter contains about 85% hydrogen, Saturn, 75%. Hence, the theory holds, the distant giant planets are largely made up of light elements, with hydrogen predominating.

True, not everything by far is clear, and there certainly remain all kinds of problems. And not only about details, but concerning the most fundamental aspects as well. Take, for instance, the "prehistory" of our planetary system. Observations suggest that the major portion of the mass of the solar system is concentrated in the sun—99. 87%—yet the angular momentum of the sun rotating on its axis comes out to only 2% of the angular momentum of the entire system. This distribution of momenta is readily obtainable from the Shmidt theory, provided the cloud which the sun captured already had

a tremendous angular momentum. The obvious question is: Did the cloud have angular momentum?

There are other questions too and in no small numbers. Shmidt's theory is not generally accepted at present. New hypotheses are constantly being advanced.

Gravity on the earth

Let us now come down to earth and talk about gravity at home. Here is how a French pilot and poet described his impressions after a forced landing in the African desert. He woke up on a sand dune looking at the stars. "At first I couldn't understand what this expanse of space was before me, and not finding anything to grasp at, no rooftops, no trees between me and these depths, I felt dizzy and detached, as if I were hurtling down an abyss.

"Yet I didn't fall at all. From head to foot I was tied to the earth. And I lay heavy on it with all the weight of my body, feeling some kind of reassurance. The force of gravity appeared to me all-powerful, like love.

"I felt the earth supporting me, pushing me up and carrying me off into the night space. I discovered that the weight of my body was pressing me to the planet like one is pressed into his seat when a car takes curves at high speed. I revelled in this support, in its solidity, its reliability and felt the curved deck of my ship under the weight of my body...

"I felt in my shoulders this force of gravity, so harmonic, constant and eternally the same. I was linked to mother Earth."

Yes, a mother, indeed it is—thanks to gravity.

Just think for a moment what role gravitational forces play in the life of our planet. Whole oceans open up, oceans of phenomena, and real oceans of water and oceans of air, too. The ocean of our atmosphere would not exist if it weren't for gravity.

The waves on the seas, the motions of every drop of water in the seas and lakes and rivers, all streams and currents and winds and clouds, and the whole climate of our planet are determined by the interplay of two basic factors: solar radiation and the earth's gravity.

Gravitation not only holds all beings and things and

water and air, but compresses them. This compression at the earth's surface is not so very great, but it plays an enormous part.

A ship at sea. It doesn't sink. Why? Because of the familiar buoyancy force of Archimedes. And this force is generated by the compression of water due to gravitation, the compressive force increasing with depth. Inside a spaceship in interplanetary space there is no buoyancy force and no weight.

The globe itself is squeezed by gravitational forces to horrendous pressures. In the centre of the earth, the pressure probably exceeds 3 million atmospheres.

Under the prolonged operation of gravitational forces, what we know as solids become fluid and behave like tar. The heavier materials move downwards towards the core of the earth and the light-weight ones come to the surface. This process has been operating for thousands of millions of years. Shmidt's theory states that it is still operating. The concentration of heavy elements in the earth's centre continues to build up slowly.

The tides

Now about the effect on the earth of the attraction of the sun and our closest celestial neighbour the moon. Only the inhabitants of ocean shores can observe these effects without special instruments.

The action of the sun is practically the same on everything on the earth and inside. The force with which the sun attracts a Muscovite at noon, when he is closest to the sun, hardly at all differs from the force acting on him at midnight. This is because the earth-sun distance is ten thousand times greater than the earth's diameter and an increase of one ten-thousandth when the earth turns half a circuit on its axis hardly at all changes the force of attraction. That is why the sun imparts almost the same accelerations to all parts of the globe and to all bodies on its surface.

Almost, yet not exactly the same. It is this slight difference that accounts for the ocean tides.

On the sunward side of the earth, the attractive force is somewhat greater than is required for motion of this portion in an elliptical orbit, while on the opposite side

of the earth it is slightly less. As a result and in accord with the laws of Newtonian mechanics, the water in the ocean is pulled up towards the sun on the sunward side and retreats from the land on the opposite side. These tidal forces stretch the globe and give the seas an ellipsoidal shape.

The smaller the distances between interacting bodies, the greater the tidal forces. That is why the shape of the world ocean is more affected by the moon than by the sun. We spoke of the sun simply because the earth revolves round it, and this simplifies explaining deformations of the ocean surface.

If the parts of the earth were not held together, tidal forces would tear our globe to pieces. This may have happened to one of Saturn's satellites when it came too close to the large mother planet. Saturn's wonderful ring is probably made up of the fragments of what was once a satellite.

Thus, the surface of the oceans is like an ellipse, the major axis of which is turned towards the moon. The earth rotates on its axis. That explains why tidal waves move

over the oceans in a direction counter to the rotation of the earth. When it approaches the shore we have flood tide. In some places the water level rises to 18 metres. Then the tidal wave recedes and we have ebb tide. The water level in the ocean fluctuates, on the average, with a period of 12 hours and 25 minutes (half a lunar day).

This simple picture is quite considerably distorted by the simultaneous tidal action of the sun, the friction of the water, the resistance of the continents, the complex configuration of the ocean shores and bottom in littoral zones, and by certain other special effects.

An important point is that the tidal wave slows down the earth's rotation. True, the effect is very small—the day grows longer by one thousandth of a second every hundred years—but operating through thousands of millions of years, the retarding forces will ultimately turn the earth towards the moon so as to face it all the time, and the earth day will become equal to the lunar month. The moon has already experienced this. It has been retarded to such an extent that the same side is always facing the earth. To see the other side, we had to send spaceships round the moon. Mercury, the planet closest to the sun, has also lost its rotation, the same side of the planet always facing the sun. One side is terribly hot, while the other, dark, side experiences cosmic cold. Apparently, that too will be the fate of our earth—only tens or hundreds of thousands of millions of years from now.

Is the gravitational constant diminishing?

Quite naturally, we do not know with certainty a great deal of the past or the future of our planet. But do we know everything about the earth even for the present time? Is everything clear, at least "gravitationally" speaking?

Indeed, the gravitation of the earth has been studied in great detail. There is even a whole science called gravimetry. Gravimetrists have studied so thoroughly all the peculiarities of local changes in gravity that they can point to large concentrations of heavy minerals deep under

the earth, or of voids and masses of particularly lightweight materials.

Comparing the true trajectory of a satellite with the computed flight path, successful attempts have been made to redefine the shape of the earth and to establish to what extent its form differs from a regular sphere. And that is not all.

One of the most astounding observations that scientists are discussing more and more is that our planet seems to be increasing in size, continuously yet very slowly. Something on the order of only a fraction of a millimetre a year (speaking of the average radius), but this is no mean figure on the geological scale. The length of day is changing, hence also the rhythm of heat and light, and consequently the climate too.

If the earth is expending, the natural question is: Why? Under the action of what forces? One might suggest the geology of the internal layers of the earth, which have not yet been studied sufficiently. Perhaps, on the other hand, we are dealing with an entirely new and specific property of gravitation, which up till now has escaped our attention. It might be that the very force of universal gravitation falls off very slowly with time.

This idea underlies a recently advanced hypothesis. The force of gravitation is *proportional*, as we have said, to the product of the masses and inversely proportional to the square of the distance between them. The law of universal gravitation must thus have a factor of proportionality. It is usually called the *gravitational constant*. Now it may be that this gravitational constant is not in the least a factor that depends solely "on the choice of units" as physicists are wont to say, but that it is slowing changing with time. If that is so, then all the theories of the origin of the solar system must be revised.

Who knows? The point is that we are still in the dark about many things concerning gravitation. Newton formulated his famous law of universal gravitation and then posed the ultimately profound question: What is gravitation, what is its nature and how is the interaction transmitted between gravitating masses?

Newton only *described* gravitation. The time has come to *explain* it.

Let us now see what has been done in this sphere.

4

Looking for an intermediary

The great Danish scientist Niels Bohr once characterized the theory of electromagnetism as a rational way out of the limits of classical mechanics, a way out "suitable for alleviating the contrast between action at a distance and action on contact".

This contrast is still more striking in the problem of universal gravitation because, for one thing, the distances are often fantastic.

Though not everyone is able to explain the complex mechanism of transferring effort from the hand along a chain to a pail of water in a deep well, one thing is clear: if a single link is cut out of the chain, the mechanism breaks down.

For a long time, gravitational forces were visualized as a remarkable chain without a single link. In science, this is known as action at a distance, with no intermediaries.

It must be noted, however, that although physicists at times got used to action at a distance and even found it convenient, they could not completely accept the fact that two bodies could attract or repulse each other through absolutely empty (or—to take the other extreme—absolutely filled) space.

The search for an intermediary for gravitational interactions began almost at the same time as these forces were hypothesized. Newton himself fully grasped the extent of this physical problem.

The surprising thing, one might think, is that having stated quantitatively the famous law of universal gravitation Newton should have so emphatically turned away from any search of a mechanism of transmission (which led his numerous commentators to place him in the camp of those supporting action at a distance). "Turned away" is of course not the proper expression. The crux of the matter is bound up with two important circumstances.

First of all, Newton was not able—due to the level of science at that time—to find an explanation of the nature of gravitation. This would have required such fundamental advances as the development of the concept of fields, which we shall discuss later, the creation of electrodynamics and, finally, the theory of relativity.

Secondly—and this is not so obvious to the present-day researcher, though it probably played quite an important role—it is a matter of the very approach to natural science, its methods and problems.

Many probably know of the struggle which began in the seventeenth century between the Cartesian and Newtonian systems of ideas in natural science. Descartes (Cartesius), Gassendi, Francis, Bacon, Hobbs, Locke and other outstanding thinkers of the time were able—and this is primarily due to Descartes—to make a decisive advance from the dominant medieval scholastic philosophy with its attempts to explain nature by the interplay of all kinds of "sympathies" and "antipathies" and with its idea of purpose in phenomena. However, important as this new school was with its characteristic union of philosophy and natural science, we must admit that the exact sciences (in the modern meaning of the term) frequently resulted from struggles against this school. Exciting as the theoretical speculations of Descartes were, they lacked one very essential thing—they were not supported by experiment, to some extent were even opposed to experiment. That is what caused Huygens to remark ironically: "Descartes apparently is ready to solve all the problems of physics without worrying about whether he is reasoning correctly or not." Characteristic in this connection is Descartes' attitude to Galileo, whom the French thinker accused of what is now termed "empirism". Descartes believed that Galileo did not consider the underlying cause of things and only studied the foundations of certain specific phenomena, thus building without a foundation.

"Hypotheses non fingo"

Newton took the path indicated by Galileo. One had to relieve science of concepts that did not stem from nature,

one had to bring to a halt the endless stream of hypotheses in the spirit of Cartesian philosophy and turn to a study of the true laws of nature. Newton believed that everything that does not stem from phenomena is hypothesis, that there is no place for hypotheses in experimental physics, and that in the latter one deduces certain propositions from observed phenomena and generalizes them by means of induction. Newton's dictum "Hypotheses non fingo"—I frame no hypotheses—should be understood precisely as a rejection of the abstract speculations of Cartesian philosophy.

Newton's extremely negative attitude towards the production of hypotheses is also evident in the question of the nature of gravitation. However, it would be quite incorrect to interpret this as acceptance of the idea of action at a distance. Incidentally, Newton himself clarified the point in a letter to Bentley, where he wrote that he considered absurd the assumption that a body some distance from another body could act on it through empty space without any intermediary agency. For that reason Newton believed that gravity should be caused by some agency constantly operating in accord with specific laws.

What the nature of this agent was remained unsettled. And no solution was given in subsequent discussions of the problem by such illustrious thinkers as Johann Bernoulli, Huygens, Leibniz, Daniel Bernoulli, Lomonosov, and Euler.

All the vogue at one time was the rather naive theory of "streams". On this theory, space in all directions is permeated with streams of matter (the nature of these streams was never specified). If you imagine two bodies separated by a small distance, they appear to shut one another off from these streams. The result is that on the outer side the streams (and, hence, the pressure) are greater than on the adjacent sides. This difference in pressure was taken to be the cause of universal gravitation. Such an explanation can hardly be called satisfactory. It not only introduces very essential hypotheses, but leads directly to conclusions that do not at all fit into any kind of experimental framework. For one thing, it predicts nonexistent gravitational shadows.

Good taste

The problem of gravitation was again discussed 234 years after Newton established the law of universal gravitation, but this time from a fundamentally different angle. To take this new step, it was necessary to reconsider the most fundamental of all concepts—space and time. Actually, to take this step forward in understanding the nature of gravitation meant constructing a new physical system of ideas. And now, looking back, one is amazed that the titanic job—and without the slightest exaggeration one can say it was a revolution in physics—could be accomplished practically by a single man. His name was Albert Einstein. Perhaps it would be no overstatement to say that not a single physical theory ever gave rise to such a tempestuous, passionate interest among the entire community of physicists (and nonphysicists) as Einstein's theory of relativity. Scientific journals and books discussed it. But in the twenties, these discussions of a purely scientific event spilled over into every newspaper and magazine including children's papers and the fashion journals. True, in all fairness it must be said that the number of those who wrote about relativity far exceeded the number of those who really understood the theory. But this very fact of tremendous interest of the mass of the people in problems of gravitation, about which just a short time before no one had even thought, is undoubtedly symptomatic. What was the reason? Particularly since Einstein's discovery—for one thing his theory of gravitation which we shall discuss later on—did not have (and hasn't even now) any practical or tangible applications. It has not helped to design a single machine nor has it helped to feed or clothe anyone, yet Einstein has been talked about, discussed and argued over more than any other scientist, who perhaps has done a great deal in satisfying the practical needs of society. It has nothing to do, of course, with passing fads or publicity, and it isn't that the new theory attracted people because of its boldness and apparent paradoxicalness. Apparently, the decisive factor here was that the theory of relativity enormously widened scientific vistas and touched on the most fundamental of all philosophical problems of natural science, posing as it did such utterly new questions as

the relationship between space and matter. As Infeld put it, humanity displayed good taste in that it gave full due to the greatness of Einstein's work in the theory of relativity.

Euclid's axioms and experimentation

Before taking up Einstein's interpretation of gravitation, let us digress a bit and arm ourselves with a few ideas which we will need later on.

We shall have to talk about geometry, more exactly about space and time. No connection with gravitation? That is where you are wrong. It was precisely the profound study of physical space and time that enabled Einstein to grasp anew the concept of gravitation. But let's not hurry.

Descartes once expressed himself in the following wonderful words: "In order to reach the TRUTH, it is necessary, once in one's life, to put everything in doubt—so far as possible." To doubt something that appears to be so obvious that no doubts are possible. To break through the magical circle of common sense, which so often seems obvious simply because no one has taken the trouble to ponder over it.

For centuries, the schools of all countries have taught geometry as a strict system of Euclidean theorems. All these theorems follow logically from ultimately simple, absolutely trustworthy starting propositions called the axioms of Euclid.

Euclidean geometry was taken into physics in toto, without any reservations and actually without the slightest hint of any necessity to check them. Both Galileo and Newton viewed space as an impassive cold background. Time flows as if governed by some sort of absolute WORLD timepiece, counting off the seconds for the entire universe. What is more, this timepiece cannot be altered by matter or the character of its motion. Up to the start of this century, that was the unshakable view of space and time.

But perhaps we can check up on Euclid's axioms. Would it be possible to test them experimentally? Here, two approaches are conceivable. There will of course

be those who will oppose such an inspection. They will say that geometry, like many other branches of mathematics, should be regarded as a purely logical construction and will therefore refuse to test its propositions against experiment. This viewpoint is fully justified in all cases, except one—when you are interested in the geometry of "real" actual physical space. That is exactly what we want at present: not some kind of abstract "mathematical" spaces, but real space. Which means that the final and decisive word belongs to experiment. This is very important, for experiment may not want to be squeezed into our familiar concepts. And then it will be necessary to reconsider much of what had seemed so obvious. Experiment, even when it takes up the study of such a "nonmaterial" object as space, reduces ultimately to the observation of matter in its multifarious forms. And this almost inevitably (later we shall see that the word "almost" can be dropped) leads to an establishment of relationships between the behaviour of matter and the nature of space. This may at first glance appear rather out of the ordinary, yet if one gets to thinking about it, the more customary notion of space (and time, we may add) as of some kind of "impassive cold background" on which events develop will then appear to us as stranger still.

And one final point. If it is experiment that we are to deal with, then it should be clear from the outset that no experiment can be absolutely exact. Mistakes (errors, to put it more delicately) in experimentation—in even the finest experiments—are inevitable. They are due to imperfections in our instruments, to accidental effects and, occasionally, to the physical essence of the phenom-

enon itself (for instance, no matter what you do, the dimensions of solids cannot be determined to less than the order of magnitude of intermolecular distances). This should ever be kept in mind in all physical investigations. We too will bear in mind the fact that no matter how rigorous our theories appear today, physical geometry for one, they are all approximate, and tomorrow or the next day they may be modified quite essentially.

Lobachevsky's geometry

That Euclidean geometry is not the only logically consistent and possible geometry was first clearly formulated last century. The first geometry with postulates different from Euclidean geometry was proposed by the great Russian mathematician Lobachevsky. Independently, the Hungarian mathematician Bolyai developed another non-Euclidean geometry. Non-Euclidean geometry was so unconventional that only three of four of Europe's greatest mathematicians could actually grasp its significance. In Russia, Lobachevsky was so improperly understood that in the obituary his administrative duties were mentioned but not a word was said about the new geometry he had constructed.

However, great ideas in science never die out even if at the time of their discovery they appear bizarre and paradoxical. More, time itself becomes one of the proofs of their invincibility. By the end of the nineteenth century, several non-Euclidean geometries had already come into existence. Most important of them for physics was Riemann's geometry. However, what was still lacking was an essential push for these purely logical constructions to "come to life" as a burning issue.

Einstein's special theory of relativity

An experiment was needed—an unusual experiment, where the object of investigation was geometry! Gauss had thought about it, privately, for his contemporaries

might easily have made him out to be a crank! His experiment was to test whether the sum of the angles of a triangle is equal to "two d". Lobachevsky had spoken of this experiment. Thus appeared an experiment that dealt the first blow at customary concepts of space and time. However, at first glance there didn't seem to be any relation to geometry.

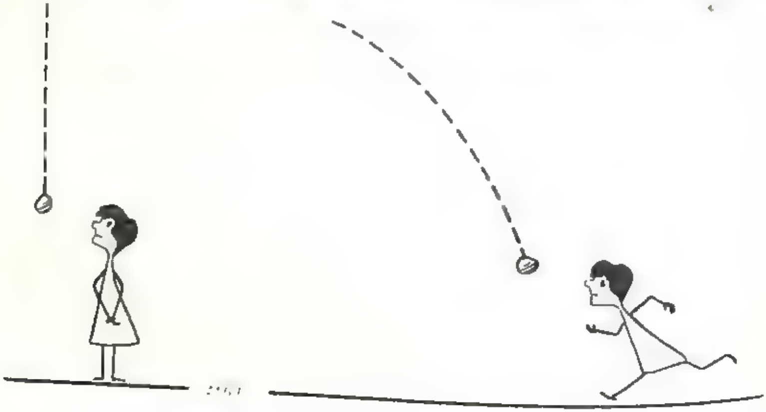
A whole series of independent and diversified experiments led to a conclusion that fascinated physicists by its paradoxicalness: no matter how an observer was moving when measuring the velocity of light, he always obtained the same numerical result. Whether you are standing still or trying to catch up with the ray of light, the velocity of light remains absolutely the same. One can't run away from a ray of light any more than he can from his own shadow. And the velocity will be the same if your instruments are coming head-on into the light.

This didn't at all fit into the framework of Galileo-Newtonian concepts, testified to the approximate nature of the latter, and indicated the urgent necessity of building a new theory that would enable one to comprehend this experimental result.

The decisive step in the construction of this new theory was taken by Einstein.

It would take us too far afield to discuss Einstein's special theory of relativity. All we need is some of its results. But first a few words about relativity as such.

Galileo already understood the relativity of mechanical motion. One cannot say simply that a body is in motion.



One has to specify with respect to what bodies it is moving (as physicists put it, with respect to what system of reference). The outer pattern of motion is, of course, different in various reference systems. The walls of a railway car are stationary with respect to the reference system of the passengers sitting there. And the very same walls are in motion relative to the reference system fixed in the earth. The flight path of a vertically falling stone looks different to a stationary observer and to a rapidly moving observer. Everything is relative: the velocity, the distance traversed by bodies and the trajectory. However, there is something that does not depend on the choice of reference system—the laws of motion themselves, Newton's laws. In all inertial systems, these laws are absolutely identical.*

What this means is that when sitting in a closed cabin, you are unable—through any kind of mechanical experiments—to determine whether the cabin is at rest or in uniform motion. Or we can put it differently: all inertial systems of reference are equivalent. One cannot find an "absolutely stationary" one nor can be find one "in absolute motion".

Einstein generalized this principle not only to mechanics but to all other processes. The experimentally confirmed fact of the constancy of the velocity of light was taken as the second initial requirement that the new theory had to satisfy.

For our discussion we shall need only one important corollary of Einstein's theory of relativity, namely, the so-called Lorentz contraction. If we measure the length of some rod in a reference system at rest and then measure it in a system with respect to which this rod is in motion (longitudinally), the latter length will be found to be less than the former. And the gist of the matter is not that something changes inside the rod. The very same may be said of two independent particles separated by a considerable distance. What changes is the geometry itself, the scale of length in the direction of motion.

Another thing to note is that the pace of a clock differs in different inertial systems. It is fastest in a system rela-

---

* To a high degree of precesion, an inertial system is a reference system the centre of which is fixed in the sun and the axes of which are directed toward fixed stars, or any other system moving with constant velocity relative to this system.

tive to which the clock is at rest. In any other system, time flows more slowly and this (like Lorentz contraction) becomes more marked as the velocity of the system approaches that of light. Incidentally, since Galileo-Newtonian mechanics was born from observations of the motion of relatively slow bodies (moving at velocities much smaller than the velocity of light, which is close to three hundred thousand kilometres per second), one could speak of a unitary "absolute" time and one could easily ignore Lorentz contraction.

Principle of equivalence

However, what connection can there be between relative lengths (Lorentz contraction) and the problem of gravitation? That precisely is our problem. Does the reader remember what we said at the beginning of this chapter about a remarkable property of gravitational forces?

All bodies, irrespective of their nature and mass receive identical accelerations due to gravitational forces. How is this to be explained? It is surely no coincidence.

Thinking about this problem, Einstein pondered on a certain circumstance which was of course familiar to all physicists, but which was never linked up with gravitation. To understand what this is about, imagine that you are in the cabin of a freely falling spaceship (with engines cut off). You are in a state of weightlessness, as if there were no gravitation at all. A pendulum will freeze in its deflected position, water spilled from a glass will hang in the air as a huge spherical drop, and alongside will be all other objects irrespective of mass and shape as if suspended by fine wires. You push a heavy weight, and it glides smoothly across the cabin. Its motion would be absolutely uniform if it weren't for the resistance of the air.

Note that the spaceship does not need to be far removed from the stars and planets so that their gravitational pull should not exert itself perceptibly. Weightlessness is a fact observed on all spaceships that circle the earth, though, of course, these craft are undoubtedly within the range of almost the same gravitational forces that act at the earth's surface. The spaceman does not feel these

forces for the following relatively simple reason. The spacecraft is in motion due to two components: uniform horizontal motion and accelerated vertical motion towards the centre of the Earth. We have already spoken of the fact that it is impossible to "notice" uniform motion judging by the behaviour of things inside the cabin. Now about falling. All the things in the cabin are really falling due to the earth's attraction. But remember that they are falling with identical acceleration. And the acceleration is the same for the floor, walls and ceiling of the cabin. The astronaut falls one metre, and so also does the chair he is sitting in. And so he can just as easily hang freely above his seat.

In other words, the gravitational forces that are evident in a reference system fixed in the earth, disappear in a freely falling system (naturally, only in the limited space of the cabin and not throughout the space about the earth).*

The word "disappear" was used on purpose, for no experiment dealing with any kind of phenomenon is able to exhibit the slightest sign of gravitation when performed in a closed cabin (or a falling lift, to use Einstein's phrase).

It is worth noting that we constantly come up against this fact even without getting into a spacecraft. Our earth too is a huge space traveller carrying along all its inhabitants and moving under the action of the sun's pull. Yet we do not feel this attraction. And it is not because the effect is small, but because the earth is simply falling round in its orbit.**

Only the tides serve as a constant reminder of the attraction of the sun and moon.

From what has been said, doesn't it seem possible to create gravitational forces in a manner similar to the way they are "removed" when passing into an accelerated system of coordinates? On the one hand, it would seem to be possible. If the mechanic of some future interstellar ship is

---

* One should not think that the walls of the cabin play some kind of role of "gravitational boundaries". Gravitation ceases to be apparent at distances where we no longer notice any change in gravitational forces either in magnitude or direction.

** The motion of spaceships round the earth and the earth about the sun differs from "simple falling" in that the motion in the latter case is rectilinear. This is seen in the "inside-cabin" experiment.

able to regulate the engine so that the speed increases 10 metres a second every second, the passengers will be in the very same gravitational conditions that we experience here on the earth. There are some doubts, however. One feels this to be a substitute of gravitation and not the real thing. Now every imitation, no matter how skilfully done, differs in some way from the real thing. But in our case one cannot find any differences. The principal property of gravitational forces is that they accelerate all bodies in exactly the same way. We might say that this property is automatically assured in an accelerated system. From the standpoint of such a system, all bodies have identical additional accelerations equal in magnitude and opposite in direction to the acceleration that the system itself has as viewed from inertial systems.

Now, with all the facts in, we make the following audacious assertion: in every sufficiently small* region of space, it is impossible by physical experimentation to distinguish between the motion of bodies due to gravitational forces and their motion in a suitably chosen accelerated system. Or, to put it more succinctly, gravitation at every point in space is equivalent to a properly chosen acceleration of the reference system. According to Einstein, this equivalence pertains not only to mechanical motions, but to all processes generally.

We have thus arrived at the famous equivalence principle of Einstein, which is one of the most profound hypotheses in the present theory, a principle—as we shall presently see—which inevitably leads to the establishment of a very intimate relationship between gravitation and geometry.

The geometry in a gravitational field cannot be Euclidean

The necessity of such a relationship is obvious from even such simple reasoning as the following: in our con-

---

* A region of space is considered small if gravitative action does not change when a body is moved within this region. On this view, then, a large hall is definitely a sufficiently small region, while the earth is not because we cannot ignore changes in gravitational forces in magnitude and direction.

ventional Euclidean plane geometry, the ratio of the length of a circumference to the diameter is pi ($\pi = 3.14...$). It can be obtained by dividing the number of very small measuring rods placed around the circumference by the number of rods that fit on the diameter. Now let us see what this ratio is equal to from the viewpoint of a reference system rotating with the circle. Suppose the experimenter, in this reference system, begins to place his measuring rods along the respective lengths. We can find the result by considering this measuring process from the standpoint of an inertial system. On the theory of relativity, every measuring rod on the circumference contracts, whereas those laid down along the diameter should not experience contraction, because their directions are perpendicular to the direction of motion. Thus, the moving experimenter will place more rods on the circumference than a stationary one would, and the same number along the diameter. That is why the ratio of the circumference to the diameter is greater than pi in a rotating system. But this is only possible if the *geometry itself has changed*, in other words, if the geometry has ceased to be Euclidean! An interesting thing is that the character of the new geometry is uniquely defined by the acceleration of the separate points of the reference system.

One more step and we'll be there. When we accepted the equivalence principle, we agreed to consider all the results obtained in accelerated systems as also occurring in inertial systems with gravitation. That being the situation, gravitation may be regarded as a step back from Euclidean geometry, "curvature of space", as we shall term it for the sake of brevity.

The most remarkable conclusion of any that physics had ever come to: gravitation is linked up with the curvature of space! The "agent" that Newton spoke of and the mysterious vortices of the Cartesians are, in actuality, simply properties of space itself, the geometry of space.

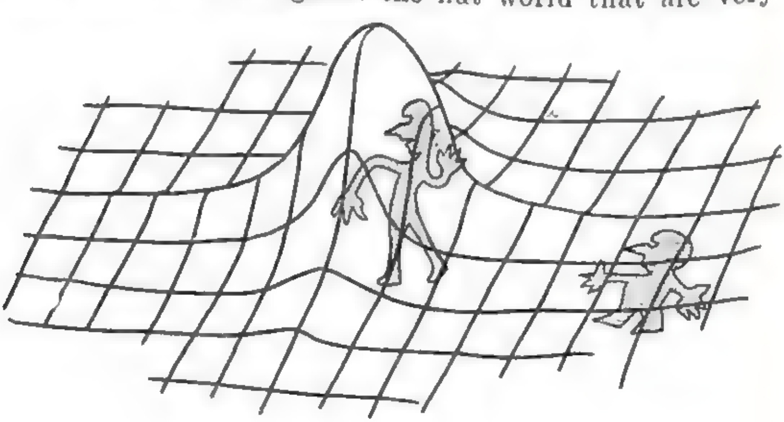
Let us take a simple model and examine this abstract and extremely complex conclusion with the aim of obtaining a graphic picture.

The geometry of two-dimensional beings

Imagine a taut rubber sheet with a surface grid. The grid will play the part of a coordinate reference frame. This is a model of space (two-dimensional, not three-dimensional) with Euclidean properties. Let us further imagine that this flat world is inhabited by some sort of two-dimensional beings with a degree of intelligence. Sooner or later, a Euclid of two dimensions will surely make his appearance. He will then formulate the principles of geometry, which will coincide with the conventional Euclidean geometry in a plane.

But suppose we press a finger into this sheet (this world) and stretch a portion of it so that the angles between lines change, and the ratio of the length of the circumference to the radius of a circle ceases to be pi, and the sum of the angles of a triangle becomes something other than $2d$—in a word, the two-dimensional geometers would definitely interpret all these events as violations of Euclidean geometry, as a curvature of space. Note also that all these events are more pronounced closer to the disturbance, which in our case is the finger pressed into the flat world.

Why not continue the analogy with the flat world? Let us compare the action of the finger pressing into the film with the action of masses that give rise to gravitation. All the more so since the finger pressure in one spot causes elastic tensions throughout the flat world that are very



much like gravitational forces (incidentally, they even fall off with distance in the same way as gravitation). However, this analogy cannot be carried too far. There is of course no elasticity involved in gravitation. The analogy is fully confined to the purely geometrical aspect of the matter.

What is a straight line?

The link-up between geometry and gravitation may be approached from yet another angle. One of Euclid's axioms reads: only one straight line can be drawn through two points. That is one of those common-sense truths, which, to follow Descartes' advice, we shall subject to some doubt. Now just what is a straight line?

Naively, we might say that it is a line drawn with a ruler. But perhaps the ruler itself is somewhat curved.

At this point the reader will probably recall that a straight line is the shortest distance between two points. Well and good, but how do we measure that distance? Again the ruler, and be sure it's a *straight* one. What we have is a vicious circle.

We might talk about "taut threads", but this would undoubtedly lead to problems of the theory of elasticity—something that is better left alone.

There is still another way—and a very simple one—to determine straight lines. Light rays are definitely straight. Let us test the ruler with a light ray and see whether it is straight or not. This is, essentially, what is done all the time in our practical life. The principle is so simple that no one ever gives it a thought.

Simple? As a practical method, yes. Yet behind this simplicity is profound physical meaning.

When using a standard, we must be certain that it is stable and is not being affected in any way by the surrounding medium. A thorough theoretical analysis indicates that a light ray is extremely stable. It does not experience any external effects.* True, that is not

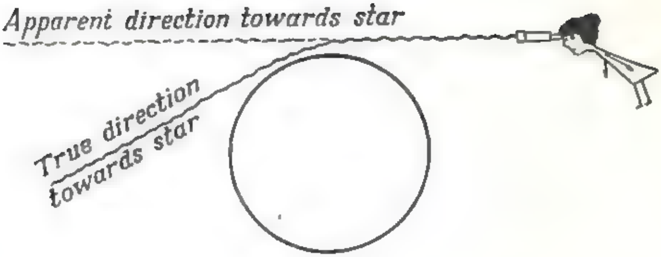
* This does not contradict the reflection and refraction of waves: these processes reduce to multiple absorptions and emissions of waves. A light wave is indeed not acted upon between emission and absorption.

exactly so. There is one force which affects light too. Remarkably, it is gravitation. The force of *universal* gravitation once again justifies its name: the effect of gravitation on light has been demonstrated by direct experiments.

The curvature of light rays

Astronomers observing the sky have been able to define the exact positions of the stars and record these positions in star atlases. They are called fixed stars and are true to their designation. A map of the sky made a hundred years ago is the same for today, to a high degree of accuracy. We have become used to this constancy. But Einstein made a marvellous prediction: during a solar eclipse, all stars located near the moon-eclipsed disk of the sun should appear to be displaced away from the sun. This displacement, or shift, was actually observed. A simple and graphic explanation is that light rays should be pulled in towards the sun under the influence of gravitation. Indeed, suppose a light ray from a distant star passes near the sun on its way to the earth (obviously, the sun's attraction will be evident only over small distances) and is deflected. An earth-bound observer who would see the star in the centre of the eyepiece if the light ray did not pass close to the sun, now sees it in the centre of the field of view if he moves the telescope slightly away from the sun. Einstein's theory gives a good quantitative and, we would like to say, natural description of the deflection of light rays due to gravitation and predicts the angles of deflection, which turn out very close to those actually measured by astronomers. We shall not go into the details of the quantitative calculations.

We shall attempt, however, to show that if Einstein's equivalence principle holds, then a light ray should inevitably be deflected. We shall start with a comparison. Suppose you are riding in a train. It is raining and the drops hit the window at an angle. If the train is in uniform motion, the tracks of the drops on the glass will be straight



lines. If the train is accelerating, the lines will be bent (curvature!). Any other streams will also be curved as judged by passengers of a train gaining speed. Streams of light are no exception.

Now recall that, in accord with Einstein's principle, acceleration is equivalent to the presence of gravitation.

Consequently, light rays (and any other rays, for that matter, streams of particles, say) will definitely be curved due to gravitation.

How an electromagnetic wave was "weighed"

The next experiment to prove the effect of gravitation on light is of a definitely mundane nature and terrestrial in scope. The reader of course knows that one has to tune his radio set in order to hear a broadcast. But he has probably never thought that the tuning might be altered by taking the radio up a mountain or down into a valley. Surely there is no reason for it changing at all. Yet it does. And here is why: first it was tuned at one level, then it was raised up and the electromagnetic waves from the transmitter had to move upwards overcoming the force of gravity in order to reach the receiving set. The experiment demonstrates that the waves will lose energy and suffer a fall in frequency. Ordinary receiving and transmitting equipment, however, is so crude that this is not noticeable. But physicists just recently were able to find transmitters and receivers with ultra-precise tuning characteristics. Don't think that these were ordinary electronic devices; in this case both transmitter

and receiver were special crystals that contained atoms with nuclei capable of emitting and absorbing electromagnetic waves of very high energy—so-called gamma quanta of very precise frequencies. In experiments of this kind carried out by Mössbauer an altitude difference of ten metres was sufficient to distinguish the acceleration* of a falling light ray. These exceedingly precise experiments are direct proof that light "weighs something" and that gravitation affects all types of matter, including electromagnetic waves.

The spectral lines of light coming from a star are slightly displaced towards the red end of the spectrum, and the shift is the greater the more massive the star. This is essentially the same Mössbauer experiment, only on a cosmic scale. It is sometimes used to measure stellar masses. This effect, just like the curvature of rays tassing near massive bodies, was predicted by Einstein.

A geometry lesson on a hypothetical planet

Now let us bring together everything that has been said. We found that the best standard of a straight line is a light ray in vacuum. Yet this very light ray is deflected, or curved, due to gravitation (and only gravitation). Thus we come to the same conclusion, only from a different standpoint. Before we spoke of "curved space", and now we use the expression "curved straight lines".

The conclusion made by Einstein about the intimate relationship between gravitation and the curvature of space was startling at the time, so unexpected and important did it seem to all dealing with the problem of gravitation. Unexpected and unusual, above all.

Recall school-day geometry. The teacher never even mentioned gravitation. He never said that only one straight line could be drawn between two points for a given magnitude of gravitational forces. True, he didn't,

---

* This "acceleration" should not be taken in a mechanical sense. There is no increase in the speed of light, which is strictly constant in vacuo (and practically so in air too), but there is an increase in energy.

that was only because Euclid's geometry grew up out of the practical activities of people living on the earth where the effect of gravitation on geometry is so slight that even today with modern instruments it is very difficult—almost impossible—to notice any imperfections in this geometry.

Now let us picture a planet (we assume that one like it exists) where the force of gravity is tens of millions of times greater than here on earth. We can visualize conditions under which a horizontal ray of light will not be able to overcome the gravitation and will bend round the planet like a satellite. And if we give free reign to our imagination, we can picture a school on this planet and a geometry teacher saying roughly the following: "Light in a vacuum moves in a straight line. Let us imagine a powerful searchlight suspended over one of the poles and sending a beam of light in a horizontal direction. We shall assume that there is no scattering, refraction or absorption. Then the light rays will move over the surface of our planet to the other pole and, bending round it, will come back to the searchlight from the other side. Turning the searchlight a bit, you will get another ray moving along a straight line and passing through both poles. And we can get any number of such straight lines. They are very much like meridians connecting the poles. Thus you see, children, that it is possible to draw any number of straight lines through two points, in our case two poles. Remember this axiom, for it is one of the basic principles of geometry. It might interest you to know, children, that mathematicians have thought up a geometry in which only one straight line passes through two points. However, this geometry can hardly have any practical application."

The pupils will learn this rule and will say that parallel lines intersect, that the sum of the angles of a triangle is not equal to two *d*, and when they graduate from school and go to work they will never encounter any geometrical paradoxes.

Many other interesting things could be told about this planet, where—in principle at least—everyone could see the back of his head without any mirrors, but it is time to return to earth. What we have learned is that what is customary need not always be unique or universal. Even geometry is no exception.

One more portion of doubt

One more point. An attentive reader will probably have noticed that in all our reasoning about light rays there has been a slight element of naive practicality. In practice, we say, a light ray is a standard for a straight line, and since it curves due to gravitation, this means that gravitation itself is associated with a curved geometry. Isn't too much emphasis placed on the word "practice"?

Continuing in this direction we can readily get caught in a maze of primitive contradictions. "Practically speaking" one might say that the piece of earth in our immediate neighbourhood is just about flat. But that does not permit us to conclude that the whole earth is for that reason flat. Or take a spoon in a glass of tea, it appears broken, yet we can easily explain this fact by light refraction at the boundary line between water and air. Is it not possible, by similar reasoning, to take into account the refraction of light rays produced by gravitation and introduce an appropriate correction?

But what does that mean—introduce a correction? We have ways of convincing ourselves that the earth is spherical and a spoon in a glass of tea is not actually broken. These ways are through experiment. Spacemen at a glance see that the earth has the shape of a sphere. Now we need a vantage point that will permit us to separate geometry from gravitation. What kind of an experiment could be thought up to demonstrate that "in actuality" straight lines are indeed straight, space is flat, and it is only light rays that become curved? The point is that we would have to find an absolute standard of straightness, and none exists.

Even that isn't the whole point. One might easily reason that there is no standard of straight lines and that on earth all rulers are actually curved and that we are not capable of straightening them out. Then would our terrestrial geometry be less like Euclidean geometry? Does the sum of the angles in a triangle change in any way due to the fact that we do not use a rectangular Cartesian coordinate grid, but a curvilinear geographical grid of parallels and meridians? Of course not. Then where is the difficulty?

Everything again boils down to the principle of equivalence. Recall the discussion about the rotating disk with which we began our story.* The essential thing is not that some kind of straight lines have become curved, but that the geometrical relationships themselves no longer hold: the ratio of the circumference to the radius is now different from that required by Euclid's geometry. And now, by virtue of the equivalence principle, this very effect should be generated by forces of universal gravitation (suitably chosen, naturally).

And the curvature of a light ray, which was used as a graphic illustration of physical geometry, was not the cause but a consequence of the curvature of the geometry.

Eddington's parable

Now let us come back to the law of universal gravitation. We have spent so much time "curving" space that we might appear to have forgotten why we began this talk. Well, we haven't. What is more, the whole foregoing discussion is essentially a new interpretation of gravitation.

This is well illustrated in a parable of the English physicist Eddington taken from his book *Space, Time and Gravitation*, which we give here with our commentaries in brackets.

"A race of flat-fish once lived in an ocean in which there were only two dimensions. It was noticed that in general fishes swam in straight lines, unless there was something obviously interfering with their free courses. This seemed a very natural behaviour. But there was a certain region where all the fish seemed to be bewitched; some passed through the region but changed the direction of their swim; others swam round and round indefinitely. One fish (almost Descartes) invented a theory of vortices; and said that there were whirlpools in that region which

* The case with the disk also enables us to demonstrate the effect of gravitation on the pace of a clock. Indeed, the farther a clock is from the centre of a rotating system, the greater its velocity, hence, the slower it ticks. On the other hand, acceleration increases with the distance from the centre. Thus, on the equivalence principle we can draw the conclusion that the stronger the gravitation near the clock, the slower the clock will tick. Accordingly, one can speak of the "curvature" of time in the same sense as the curvature of space.

carried everything round in curves. By-and-by a far better theory was proposed (Newton's); it was said that the fishes were all attracted towards a particularly large fish—a sun-fish—which was lying asleep in the middle of the region; and that was what caused the deviation of their paths. The theory might not have sounded particularly plausible at first; but it was confirmed with marvellous exactitude by all kinds of experimental tests. All fish were found to possess this attractive power in proportion to their sizes; the law of attraction (like the law of universal gravitation) was extremely simple, and yet it was found to explain all the motions with an accuracy never approached before in any scientific investigations. Some fish grumbled that they did not see how there could be such an influence at a distance; but it was generally agreed that the influence was communicated through the ocean and might be better understood when more was known about the nature of water. Accordingly, nearly every fish who wanted to explain the attraction started by proposing some kind of mechanism for transmitting it through the water.

But there was one fish who thought of quite another plan. He was impressed by the fact that whether the fish were big or little they always took the same course, although it would naturally take a bigger force to deflect the bigger fish. (The sun-fish imparted identical accelerations to all bodies.) He therefore concentrated attention on the courses rather than on the forces. And then he arrived at a striking explanation of the whole thing. There was a mound in the world round about where the sun-fish lay. Flat-fish could not appreciate it directly because they were two-dimensional; but whenever a fish went swimming over the slopes of the mound, although he did his best to swim straight on, he got turned round a bit. (If a traveller goes over the left slope of a mountain, he must consciously keep bearing away to the left if he wishes to keep to his original direction relative to the points of the compass.) This was the secret of the mysterious attraction, or bending of the paths, which was experienced in the region.

The parable is not perfect, because it refers to a hummock in space alone, whereas we have to deal with hummocks in space-time. (We cannot go into this problem in

more detail in our book.) But it illustrates how a curvature of the world we live in may give an illusion of attractive force, and indeed can only be discovered through some such effect."

In a word, this may be formulated as follows. Since gravitation curves the paths of all bodies in the same way, we can regard gravitation as a curvature of space-time. Gravitation is the *alter ego* of geometrical curvature.

Nothing more need be related to the curvature of space-time than the curvature of space-time paths (which are called world lines)* of all bodies without exception.

The movement of the perihelion of Mercury

After quite some discussion we have arrived at an essentially new understanding of gravitation. It is important and interesting in itself. But perhaps it adds nothing significantly new leaving us, ultimately, with the same good old Newtonian law. Not in the least. It is not only a point of a new setting for old truths but of fundamental generalizations and new effects.

We have already mentioned the fact that Einstein's theory gives a correct quantitative description of the deflection of light rays due to gravitation, and we have spoken of the Mössbauer effect. We can also add the successful explanation of the movement of the perihelia of the planets, particularly that of Mercury. We have in view the following. Calculations based on Newtonian mechanics lead to the conclusion that the orbits of all the planets should be ellipses with unchanging positions in space. However, observations show that these orbits are slowly turning. This is most noticeable in the case of Mercury, the planet closest to the sun and experiencing for this reason the greatest gravitational effects. Calculations based on the Einstein theory of gravitation yield a good quantitative description of this phenomenon.

* A world line is a curve that depicts the dependence of the coordinates of a moving point upon time. For simple one-dimensional motion (motion along a space curve), a world line represents the sole coordinate as a function of time. A world line is straight for uniform motion and curved for accelerated motion.

Gravitational waves

Everything that we have mentioned above is rather in the nature of small corrections. Yet there is something fundamentally new that follows from Einstein's interpretation of gravitation. First of all, we must point out the conclusion relating to the finiteness of the velocity of propagation of gravitation.

In Newton's law of universal gravitation, nothing was said of the time of transmission of interaction. The implication was that it is propagated instantaneously, no matter what the distances between the interacting bodies. This view was generally typical of the adherents of action at a distance. From Einstein's theory it follows that gravitation is transmitted from body to body with exactly the velocity of a light signal. If some body is moved, then the curvature of space and time generated by this action does not change instantaneously. At first the effects will be apparent in the immediate vicinity of the body and then extend out to more distant regions of space, and finally, in the whole region of space, there will be a "new distribution of curvature" that corresponds to the altered position of the body.

We have now arrived at a problem that is still causing much argument and clashes of opinion—the problem of gravitational radiation.

Can gravitation exist without a mass generating it? According to the Newtonian law, the answer is an unconditional NO. It would be senseless to even pose such a question. However, as soon as we agree that gravitational signals are propagated at a finite velocity (large though it may be), then everything is radically changed. Indeed, suppose a mass (a ball, for instance), which gives rise to gravitation, is at rest. All bodies about the ball will be affected by conventional Newtonian forces. Now let us remove this ball from its original position with a large velocity. For a moment, the surrounding bodies will not "feel" the difference, for gravitational forces do not change instantaneously. Time is required for the changes in curvature of space to spread out in all directions. Which means that the surrounding bodies will, for a time, be experiencing the former

action of the ball after the ball has left its original position.

Thus, the curvature of space (more precisely, the curvature of space-time) acquires a certain definite independence such that a body can be extracted from a region of space where it has generated curvatures and in such a manner that the curvatures will remain (at large distances, at any rate) and will develop in accord with their own laws. Such is gravitation without a gravitating mass! We can go even further. Suppose we make our ball oscillate. Then, as follows from Einstein's theory, ripples or waves of gravitation will be superimposed on the Newtonian picture of gravitation. To get a better picture of these waves, let us revert to our rubber-sheet model. Press your finger into the rubber and also move it back and forth. You will see that these oscillations will be transmitted in all directions over the taut sheet. This is an analogy of gravitational waves. The farther away from the source, the weaker the waves.

Now remove your finger from the rubber sheet. The waves do not disappear. They continue to exist independently, spreading out over the sheet and generating, in their path, a curvature of the geometry.

In exactly the same way, the "waves of space curvature"—gravitational waves—can have a separate existence. Many scientists have drawn this conclusion from Einstein's theory. However, they go further. They say that gravitational waves are not only emitted, but can also be absorbed, like any other type of wave. The effects are very feeble, though. For instance, the energy released by a burning match is many times greater than the energy of the gravitational waves emitted by our entire solar system during the same time. The important thing here, however, is not the quantity but the principle of the effect.

Do gravitational waves exist?

As soon as we say the "energy of gravitational waves", we have to ask about the carrier of this energy. Obviously, as the name indicates, the carrier is the wave. But

we have been speaking of "waves of space curvature". Can the curvature of space have energy? This is no simple bending of an elastic rod, remember. There, a certain amount of energy has to be expended to produce the deformation. Our "curvature" is simply a deviation from Euclidean space. Note that the space is empty.

The situation gets more complicated due to the fact that gravitation can be "removed" from any point of space and we go over to the appropriate "falling-lift" reference frame mentioned earlier. Can we then say that "curved space" has a store of energy, when one system has curvature and the other has none?

"Yes," say the gravitational wave supporters. This is because the "falling-lift" system straightens out space and by no means switches off gravitation at once and everywhere. If gravitation is removed at one spot, there is a build up in other spots, and vice versa. It is not possible to select a reference system moving with acceleration that would be equivalent, say, to the entire distribution of terrestrial gravitation. If a "falling lift" is located over the north pole, this means that the force of gravity ceases to be felt precisely over this pole, but in this same system the attractive forces over the south pole should double. That is precisely the case in a gravitational wave, the choice of the reference frame only affects the redistribution of energy in space, but does not in the least mean that we can obtain, at will, either a zero energy or some energy different from zero.

The followers of gravitational waves (and they seem to be in the majority now) predict yet another fantastic thing: the conversion of gravitation into such particles as electrons and positrons (they always come in pairs), protons and antiprotons, and so forth.

The picture is something like this. A gravitational wave reaches some portion of space; at a certain instant this gravitation undergoes a radical and sudden change, diminishing and simultaneously giving rise to, say, an electron-positron pair. This may also be described as an abrupt, saltatory decrease in the curvature of space with a simultaneous generation of a pair.

Many attempts have been made to translate this into the language of quantum mechanics. Particles called gra-

vitons have been introduced to correlate with the non-quantum image of the gravitational wave. In the literature we find the phrase "transmutation of gravitons into other particles". These transmutations (mutual transformations) are possible between gravitons and, in principle, any other particles. This is because there are no particles that are insensitive to gravitation.

These transformations may have a very low probability and occur very infrequently, but on the cosmic scale they may be fundamental.

So far we do not know whether such things occur or not.

Fresh opportunities

It is curious to note that the possibility is being actively discussed of harnessing these gravitational waves, of perhaps making "gravio-receivers" and "gravio-transmitters". This type of communication would have a tremendous advantage over others. Gravitational waves pass through all matter with hardly any absorption, whereas electromagnetic waves hardly at all penetrate electrically conducting media (sea water and the earth included). The problems here are legion, however. Before thinking of sending a "graviogram", we have yet to demonstrate experimentally that gravitation has a finite velocity of propagation. The effects are too small, gravitational forces are too weak. But work goes on intensively both here and abroad.

The finite, yet infinite universe

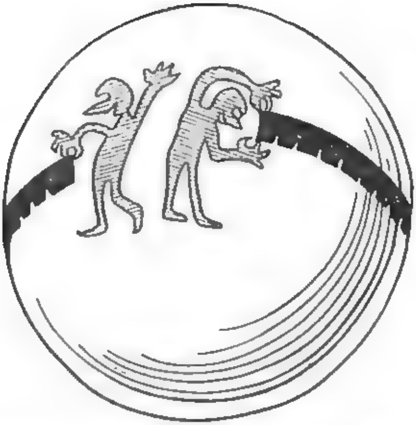
When discussing gravitational transmutations of particles we touched on cosmology. Here again it must be said that Einstein's theory was very stimulating for the development of cosmological conceptions. It gave rise to a spate of new theories and ideas and reshaped completely that oldest of all the sciences of nature—cosmology.

What do we know about the universe? Especially since our instruments permit us to look over only a tiny corner. Numberless myriads of stars are so far away from the

earth that our optical and radio telescopes cannot reach them. But the investigator mentally reaches to still farther distances. And what lies beyond? Conjectures based on general considerations that everything is the same out to infinity were somewhat satisfying but did not spur the imagination. Then in 1917 a hurricane swept the scientific world in the form of Einstein's theory of a finite universe. Finite? Then there must be a boundary. And what lies beyond the boundary? Are we to think that space ends?

Such were the questions that immediately confronted physicists. Let us first try to figure out what we mean when we say finite space. Here again we shall make use of our flat rubber-sheet world in two dimensions that we used to explain the curvature of space. We tacitly assumed that our sheet extended without bounds in all directions, and only in places where matter is found does it curve (bodies generate gravitation). In our model, the world was infinite. But what if the matter is distributed in a more or less uniform fashion? Then the curvature should be more or less the same at all places. How do we visualize a sheet with the same curvature everywhere? Nothing could be easier. A toy balloon is the best example.

Now take a look at the two-dimensional flat investigators working on the rubber sheet. For them, the surface of the sphere is their whole space. Indeed, if they sent out an expedition with the object of moving all the time in the same direction (in a straight line; that is, one which curves all the time) the expedition would, to the wonder of its organizers and participants, return to the point from which it started—only from the other side. More expeditions could be sent out, but they too would all make a complete circuit and return to the point of departure. To what conclusion would the flat men of science come? That our universe is by no means infinite, it has finite dimensions. At the same time it is limitless, since no expedition ever found any kind of a boundary. Limitless but not infinite. Those are the precise words that Einstein used to describe our space in his theory. Their meaning is, on the whole, the same as for our model: if we picture a spaceship flying in a straight line, it will finally return to where it started from (if it does



not collide with any celestial bodies, naturally). One can replace a spacecraft with light, the fastest of all travellers. Moving in gravitationally curved space, it will close in on itself eventually, after having traversed the vast expanses of limitless, yet finite, space.

An expanding universe

However, this is not all. Einstein's theory of gravitation not only enables us to speak of the finiteness of the universe, but also leads to a still more remarkable conclusion—that of an expanding universe.

The conclusion was first made by the Soviet physicist A. Friedman. Einstein himself was at first rather sceptical.

The underlying principle of Friedman's theory was that the universe as a whole is homogeneous and isotropic. Which means that no larger section of the universe differs in any of its properties from the others. All directions in the universe are absolutely equivalent. The mean density of matter is everywhere the same.

On this assumption, Einstein's equations of gravitation unambiguously lead to the conclusion that the universe cannot be stationary. It is expanding, so that all stellar clusters—galaxies—are racing away from each other.

Let us recall our two-dimensional model of the finite universe, our toy balloon. This balloon, the universe, is

constantly puffing out, so that the distances between any of its points are increasing. What is more, the increase is the faster, the farther the points are away from each other. This is because every centimetre of a segment connecting any points increases. The flat astronomers could surely see that.

The spectrum of a receding star is displaced in the direction of longer wavelengths. All the lines become redder. This is known as the red shift and is caused by the Doppler effect. The greater the velocity of recession, the bigger the shift.

The most remarkable thing is that our terrestrial astronomers were able to detect this phenomenon. The American astronomer Hubble found that all stellar islands of the universe, the galaxies, are receding from our own. And the farther away a galaxy is, the greater the displacement of its spectral lines; hence, the greater its relative velocity of recession. This velocity, u, satisfies a simple law: $u = Hr$, where r is the distance to the galaxy, and H is Hubble's constant.

The most distant galaxies are racing away from our system with fantastic velocities of the order of 100,000 km/sec, which is about one third that of light itself. In this case, the blue-green light of such galaxies is seen as red light.

Hubble's law follows directly from Friedman's theory. H falls off in inverse proportion to the time, and, consequently, the rate of expansion of the universe should slow down.

This amazing theoretical prediction merged with a marvellous experimental discovery. No wonder that the scientific community—and even the ordinary world—was struck by the audacity and originality of the cosmogonic ideas of Einstein and Friedman. It was in fact hardly short of a revolution.

The past and future of the universe

The fact that the universe is expanding (to be more precise, we should say "our portion of the universe") is quite definite. It is an experimental fact. And theory

supports it. But what will happen to the universe in the future? What was it like in the past? And finally, is the universe really finite or infinite?

We have no definite answers to all these questions, but much can be said even now, if we are ready to accept the assumption of a homogeneous and isotropic universe.

Let us first talk about the future of the universe. Strange as it may seem, this is most definite in our minds. There are only two alternatives.

According to theory, everything depends on the relationship between the mean density of the universe at the present time ρ and a certain critical density $\rho_c \frac{3}{8\pi} = \frac{H^2}{\chi}$, where H is Hubble's constant at the present time, and χ is the gravitational constant.*

If ρ is less than ρ_c, then the universe will never cease to expand. The rate of recession of the galaxies will gradually decrease, but expansion will never give way to contraction. The galaxies will move away to incredible distances, and our stellar island will become utterly lost in the vast void of space.

But if ρ is greater than ρ_c, then in time the expansion of the universe will give way to a contraction, and instead of the red shift we will have a violet shift, though it is hard to predict when this will occur, if ever.

And so to learn about the future of the universe, we must know the mean density of matter in it. The density $\rho_c = 2 \times 10^{-29}$ g/cm^3 is known because Hubble's constant and the gravitation constant can be measured with sufficient accuracy.

The main difficulty is to measure ρ. One has to know the mass of matter (which includes matter proper and radiation) not only in the stars but in all interstellar space of the visible portion of the universe. The estimates presently available are extremely contradictory. According to some, ρ is less than ρ_c, others put ρ as the greater. No definitive conclusions have as yet been obtained.

Determining the density of matter is important in yet another respect. The ρ/ρ_c ratio, upon which the future of

---

* Actually, the relationship between ρ and ρ_c is determined by the relation of potential and kinetic energy of the universe.

the universe depends, is also a determining factor for the spatial structure of the universe as a whole. For ρ greater than ρ_c, the mean curvature of the world is positive and the universe is finite. But if ρ is less than ρ_c, the universe is infinite. Thus, Einstein's theory of gravitation only shows that our former certitude about the universe being infinite may not be true, and does not assert unconditionally that the world is closed within itself.

Let us now see what there is to say about the past of the world. At one time, the universe must have been compressed into a very small volume. At that time, the density of matter was infinitely great. If we take that instant as the starting point of time ($t = 0$), then, knowing Hubble's constant, we can estimate the time of expansion of the universe. It comes out rather small: only about 10,000 million years. Which is not a lot by astronomical standards. Incidentally, the same order of magnitude for time is given by estimates based on radioactive determinations of the age of minerals.

What was the state of matter in the universe at that time? How did our universe emerge from this ultradense substance and become stars and stellar clusters? And, finally, what was the universe like before that time?

No answers are forthcoming at present, but certain possible solutions seem to be in the offing.

Obviously, gravitation was not so essential in the initial state of the universe when matter was in an ultradense state. Other forces must have played no less a role. So only after we discuss these forces can we examine the existing hypotheses.

Condensation of matter as the universe expanded into stars and galaxies was probably due to gravitation along the lines we spoke of when discussing Shmidt's theory. Later on in stellar evolution, nuclear and other forces began to play a part along with gravitational forces. We shall come back to this again later on.

Now what was the universe like prior to expansion? If we presume that the universe will continue to expand without limit, there is no answer at all.*

---

* True, just recently Zeldovich advanced a hypothesis of words being born in collisions of super-gigantic energy particles wandering about in the boundless expanses of the universe. This

It is simpler in principle to answer the question for a finite universe. For in that case, expansion should definitely be followed by contraction, and the state of the universe at $t = 0$ must be regarded as the result of a previous contraction. Then we arrive at the hypothesis of a pulsating universe. The universe pulsates (expanding and contracting) eternally with a period that is still unknown.** That appears to be the simplest and most reasonable picture of the evolution of a universe that exists eternally. Of course, in such cases the criterion of simplicity cannot be decisive.

A big step forward in the cognition of nature

One might feel inclined to ask the question: If Einstein's theory of gravitation is so complicated, if the cosmological conclusions from it are to a large extent preliminary and often of a fantastic nature, and finally if this theory has not helped to construct a single machine and has done nothing for technology as such, then what is it that compels the greatest thinkers of our day to regard it as "the most remarkable achievement of the human mind"? What is it there that has engaged the undivided attention of physicists, philosophers, astronomers and large numbers of thinking people for over forty years? It cannot be just a matter of "good taste" or the magnificent beauty of the basic principles of the theory.

The point is that this was a discovery of unusually new and involved physical relationships which students of nature had never even suspected.

With the help of Einstein's theory, we have entered a new field of unparalleled interest to humanity. Here, for the first time in an exact physical theory we have

---

way out of the impasse does not only verge on the fantastic, but is far from attractive for the physicist. Starting on this basis, it is hardly possible to investigate the distant past of the universe. Such particles should be regarded as absolutely the first in all time.

** Possibly, during pulsations, the universe does not contract to infinite density. It might very well be that the universe never, at any time, had an infinitely large density. Such is the opinion of I. Khalatnikov and Ye. Lifshits.

come closer to an understanding of the infinite. Man for the first time began to feel the pulse of the universe not in the form of poetical revelations but via precise knowledge. Without Einstein's theory, most of the questions just discussed could not even have been posed.

About a hundred years ago, Faraday was in raptures when he detected a relationship between light and magnetic phenomena. He wrote that he was able to magnetize light and illuminate the magnetic line of force.

Einstein's theory definitely cast new light on the scientific understanding of the world and the "magnetism" of its conceptions has dominated the scientific world ever since.

No matter what heights our subsequent understanding of gravitation may reach, the genius of this "Newton of the twentieth century" and the incomparable audacity with which he refashioned our picture of the world will ever remain a great step forward in the cognition of nature.

C H A P T E R   T H R E E

I sing the body electric

Leaves of Grass (Walt Whitman)

ELECTROMAGNETIC FORCES

1

The child and the scientist

A book lying on the table will never fall through the table, despite the attraction of the earth. The book will not even slip off the table if we tilt it just a little. No one is ever surprised if a person increases his speed when chased by a ferocious dog. Finally, few people give much thought to the reasons of a book, table or other solid retains its shape.

These are common facts that we encounter all the time and have been familiar with since early childhood. They have become so obvious that we never feel any need to try to account for them. In most cases, none of this is necessary in any way to get along in life. It is always important to know what is happening, but we don't always need to know why things occur one way and not another.

When still a small child, one thinks about such customary things, but they are far beyond the understanding of a child, and grown-ups rarely come back to them. The desire to explain the behaviour of "simple" things is, as the English physicist Perry put it, hidden deep down in our consciousness, and the human mind is attracted by the unusual, the unexpected. Only children and scientists ever take an interest in the commonplace.

For that reason, it is in many ways easier to talk about the unusual properties of space and time hidden in the theory of relativity than to explain why a stone retains its shape. In the former case, we immediately become interested, while in the latter is so trite as to bore us from the start.

Indeed, it is difficult to find answers to the common things we have just asked about. Attempts to find answers will take us far afield, so far in fact that we come to the very frontier of present-day science. Without going too deep, let us take a look at a series of questions which inevitably occur to anyone curious about commonplace things. So common that one feels embarrassed to term them physical phenomena.

Elastic forces and their relatives

This book is lying in front of you on the table, it is acted upon by the force of gravity. Yet it does not fall. Why? One not familiar with scientific ways will probably say "the table prevents it from falling". But this is no explanation, it is simply a statement of fact.

A person a little acquainted with physics will go farther and say that the table acts with a force that balances that of gravity. This force is called the force of elasticity and comes into play as the table sags imperceptibly under the weight of the book. But why an elastic force is generated when the table sags is a still more difficult question.

Let us stop the questions here. We shall return to the reasons for elastic forces a little later. The point is that these forces have a common origin with many other forces. They are all related and are quite distinct from the forces of universal gravitation, which to this day appear to be absolutely unique.

The forces of elasticity, which permit solids to retain their shape and oppose changes in the volumes of liquids and the compression of gases; the forces of friction, which act in opposition to the movement of solids, liquids and gases; and, finally, the forces of our muscles—these are all members of a single vast family. They are of a common nature and have a common origin: they are *electromagnetic forces.* These forces are extremely active in the multifarious domains of nature. In everyday life, with the exception of the attraction of the earth and the tides, we encounter only different types of electromagnetic interactions, if one disregards the nuclear forces of recent discovery. Even the elastic force of steam is of an electromagnetic nature. So, actually, when the steam age was superseded by the age of electricity, that only meant that we had learned to control electromagnetic forces to suit our needs.

Electromagnetic forces permit you to see this book, because light is a form of electromagnetic interaction. Life itself would be unthinkable without these forces. Living beings, man included, are able to exist in a weightless state for long periods of time, as our space ventures

have amply proven. Yet if the action of electromagnetic forces stopped for just an instant, life as we know it would cease.

In particle interactions in compact systems—from atomic nuclei to cosmic bodies—electromagnetic forces play a very essential role, whereas nuclear and gravitational forces are significant only in extremely small assemblies or on the cosmic scale. Extranuclear atomic structures, the adhesion of atoms in molecules and larger bodies are determined solely by electromagnetic forces. There is hardly a phenomenon that is not connected with the action of electromagnetic forces. Accordingly, it is difficult to even name off its multiplicity of manifestations. So far we have only mentioned a few. Where do we begin?

From what has been said, it is obviously best not to begin a study of this thing by examining one of its members, say the forces of elasticity. Then where should we begin? We have listed a number of diverse forces and called them all electromagnetic forces. But that does not mean we have explained anything.* All the more so since we usually think of electric and magnetic forces as something quite different. The force of interaction between electrified bodies is termed the force of electrical attraction and repulsion. For example, the force which makes tiny pieces of paper stick to a comb that has gone through one's hair a few times. By magnetic force we usually mean that which a magnet exerts on a current-carrying conductor, or the interaction of magnets.

So far we have only said that these multifarious interactions are of a common nature. The first question is: What is it that they have in common? Put otherwise: What forces are called electromagnetic forces?

The answer in brief is: underlying all the aforementioned forces are certain general laws called the laws of interaction of electrically charged bodies. This may sound rather repetitious but there doesn't seem to be any other way out. In the final analysis, all these forces are due to the interaction of elementary particles car-

---

* The authority behind the very term "electricity" is so great that people are frequently satisfied when they hear: "This is due to electricity." Yet that is only the beginning of an explanation.

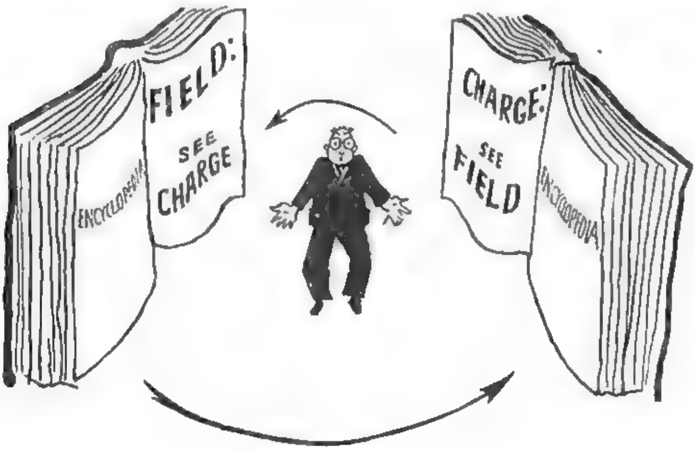
rying electric charges. Interaction between charged particles is accomplished by means of an electromagnetic field. Hence the name electromagnetic forces. If—by magic—all electric charges suddenly disappeared, there would be no forces of elasticity, friction and so forth. All bodies would break down into their component parts, even atoms would fall apart.

Our aim now is to learn about the basic laws of electromagnetic interaction. Then it will be clear why these forces are so widespread in nature and manifest themselves in such a multiplicity of ways.

2

Difficulties with definitions

What is an electric charge? We open the encyclopedia and read: "An electric charge is the property of certain particles (electrons, protons, positrons, and certain types



of mesons) wherein they are always associated with an electric (electromagnetic) field and experience specific effects of external electromagnetic fields."

Now what is an electromagnetic field? We open the encyclopedia and read: "An electromagnetic field is a physical field of moving electric charges that accomplishes interaction between the charges." What we get is a circle in which the snake is biting its own tail. A charge is that which is associated with an electromagnetic field, and a field is that which is associated with a charge.

All this is no fault of the editors. The obvious insufficiency of these definitions is simply a reflection of an actually existing difficulty of giving a reasonable definition of these fundamental notions. The point is that there is, in general, no way of giving brief and thoroughly satisfactory definitions here. That is the most important thing at present. We are used to finding sensible explanations for extremely complicated formations and processes like, say, the atom, thermal diffusion, a nuclear chain reaction and so forth. Actually, such complex structures as the atom, for instance, are not so difficult to account for. The real difficulty lies in giving a succinct and satisfying description of the very basic, fundamental concepts that cannot be broken down further into any simpler elements, which—for the present level of science—do not have any internal mechanism. This is particularly so if they are not directly perceived by our sense organs. The electric charge and the electromagnetic field are just such fundamental concepts. In school, the usual procedure is: at first they are simply not understood, then later everyone gets accustomed to them and uses them, paying no attention to their profound content.

The situation here is so involved that just recently, in the middle of last century, outstanding thinkers were capable of entertaining the most fantastic notions about electricity. Hegel, for instance, thought that electricity was "the fury proper, the raging of a body", its "own angry essence" that "is apparent in every body when it is disturbed" (from the *Philosophy of Nature*).

Electric charges and elementary particles

Let us first try to discover not what an electric charge is, but what lies behind the assertion: *a given body or particle has an electric charge.* This is practically the same thing, though not quite; the latter is somewhat easier to grasp.

At the present time, it is no secret that all bodies of nature are composed of minute, indivisible (so far as we know at present, at least) particles, which are therefore called elementary or fundamental. There is no need here to list all the particles which have been discovered to date. The important thing is that the principal role as building blocks of the world is played by electrons, protons, and neutrons. How do these particles differ?

When we say that certain particles are different, we mean that they act on the surrounding world and are acted upon by it in unlike manner. To start with, all particles have mass and their masses differ. The proton mass is 1846 times that of the electron mass, the neutron is somewhat more massive than the proton, and so forth. Accordingly, on the one hand, these particles behave differently under the action of external forces, because their inertial properties are different, and, on the other hand, the forces of gravitational interaction with each other and with the outer world differ under otherwise equal conditions.

When we say that electrons and protons are electrically charged, this means that they are capable of specific types of interaction (electromagnetic) and nothing more. The absence of a charge signifies that the particle does not have such interactions. The charge itself is a quantitative measure of the capacity of a body for electromagnetic interactions, in the same way that the gravitational mass is a quantity that defines the intensity of gravitational interactions. The electric charge is the second (after the mass) most important characteristic of any elementary particle that determines its behaviour in the surrounding world.

There is nothing exceptional in what has just been said. Even people, if one disregards external appearances, differ mainly in their reaction to the world about them and its reciprocal reaction on them. Thus, when we say that Mr. Pickwick was a jolly fellow even capable

of getting a widow out of a debtor's jail, a widow that tried to get him to marry her, we get a rather clear picture of his conduct with the people he comes into contact with. If the behaviour of a person exhibits a reciprocal relation, we get a villain like Mr. Carker of *Dombey and Son*. Then there is Pimen from Boris Godunov who feels neither pity nor anger.

In nature we have particles with charges of opposite sign. The charge of the proton is called positive, that of the electron, negative. The positive sign of the charge that a particle may have is, of course, no mark of excellence or merit. The two signs simply signify that the charged particles can either attract each other or experience mutual repulsion. If the signs of the charges are the same, they repulse, if the signs are different, they attract each other.

Just like people have more qualities than kindness and weight, so the elementary particles have properties other than charge and mass. But there is one important thing: no matter how different the properties of elementary particles may be in other respects, the charge, if there is one, is the same for all, whether we are dealing with electrons, protons, positrons, antiprotons, light, heavy and superheavy mesons. The only difference lies in the sign. There is no charge less than the electron charge. Kindness, on the other hand, like other moral qualities, is very unevenly distributed. Between an angel and the devil lies an enormous spectrum of characters.

As experiment demonstrates, electric charge in nature is conserved. The sum of the charges of all particles (taking into account the sign of the charges) remains unchanged. If a new charged particle appears (and this is a frequent occurrence) we always see the birth of another particle of opposite charge. And particles always die in pairs of opposite charges.

Charge and the laws of electromagnetic interactions

The presence of electric charge on particles presupposes *strictly definite laws* of force interactions between them. These laws are amenable to rigorous mathematical

formulation and determine the motions of the particles themselves. It is quite obvious that we know nothing essential about the charge if we do not know the laws of these interactions. A knowledge of these laws should be a definite ingredient of our general notion about the charge. Analogously, we would know nothing about a kind person if we didn't know what sort of actions there were. These are by no means simple laws that can be stated in a few words. Hundreds of volumes have been written about electromagnetic interactions and hundreds more will be written. To understand what an electric charge is, one of course does not need to read a hundred volumes, but it is necessary to learn the essentials of electrodynamics (the science of electromagnetic interactions).

By now the reader will probably have realized that it is not so easy to say what an electric charge is as it is to describe an electric train. Remember the Cheshire cat in *Alice in Wonderland*? First the grin appeared and then the rest of the cat. And it disappeared in the reverse order, starting from the tip of the tail and ending with the grin, which remained a short time after everything else had vanished. Alice had seen cats without grins, but never a grin without a cat!

Quite the same in the atomic world. There are particles with no charge, but never a charge without a particle. On present-day views, an electric charge cannot be regarded as some sort of additional mechanism that particles may have. A mechanism that can be removed from a particle and broken down into some kind of constituent parts and then assembled together again. The charge of a particle is intimately bound up with its structure, which we still know practically nothing about, in just the same way that kindness is interwoven into the whole psychic make-up of a human being. Just as there is no mechanism responsible for kind actions, so there is no mechanism that handles the "electromagnetic affairs" of a particle.

*The charge is not a mechanism in a particle, but the capability of the particle, on the whole, to interact with other particles in a specific way.**

That is the view taken by contemporary science. Don't think that our information about the charge is exhaustive in any sense and that nothing new will be added. Elementary-particle physicists are already asking the question: Why are only certain elementary particles charged and others not? Why isn't there a charge greater or less than that of the electron? How is the magnitude of the charge connected with other universal constants such as the velocity of light, the Planck constant, and so forth? Who knows, perhaps answers will be forthcoming very soon, for advances are being made all the time. Bombarding protons with very high-energy electrons has enabled scientists to establish the approximate nature of distribution of electric charge inside these particles. It was found that the proton charge is smeared over a finite portion of space (of radius close to 0.8×10^{-13} cm) and is distributed in this region very unevenly. In the centre is a compacted portion—the core—which is roughly four times smaller than the proton itself. At the same time it was discovered that there are also charged regions inside the neutron.

The most amazing thing is that despite the smearedness of the charge in space, not a single grain of it can

---

* Note, however, that an electric charge operates in exactly the same way in all particles. Other properties have no effect on the "electromagnetic behaviour" of a particle.

be torn off. Probably the most difficult thing to grasp in everything concerned with electric charge is that it does not exist in quantities smaller than a definite one.

Note again that so far we have been speaking only of the charges of elementary particles. A large-sized (macroscopic) body, as can easily be imagined, will be electrically charged if it contains an excess of elementary particles of one sign. The negative charge of a body is due to an excess of electrons over protons, the positive charge, to a deficiency of electrons. Most bodies are electrically neutral since the number of electrons in them equals the number of protons. Is the world as a whole neutral? If the universe is finite, then its electric charge is zero. In the case of an infinite universe, the total charge can be different from zero.

An essential point to remember is that electrical neutrality does not in the least signify the absence of electromagnetic properties. They are always there in latent form. Even the neutral elementary particle neutron is not devoid of them. The neutron is like a small magnet as far as its electromagnetic properties go.

3

The first steps

We will never know who it was who first noticed the marvellous property of amber, when rubbed against wool, to attract light-weight objects without actually coming in contact with them. That was long ago. According to the ancient Greek philosopher Phales of Miletus, who lived in the sixth century B. C., they were weavers.

Later it was found that glass, ebony and other substances when rubbed against fur or leather have the same

property. The Greek for amber is electron, and for this reason bodies put into this state were called electrified.

The term electricity is thus of a rather poetical origin.

These first simple experiments were the first clear demonstration of electric forces. But over two thousand years passed before a systematic study of electricity was begun and the law of interaction of electrified bodies was discovered. The unusual property of amber and certain other objects was simply an oddity, for how indeed can bodies attract each other without contact? There was nothing here to suggest that in this simple form was the manifestation of laws governing the flow of most processes on earth.

For centuries, no serious attempts were made to explain experiments with electrified bodies in a scientific manner. Suggestions that amber might have a living soul were hardly an explanation. Such experiments were carried on for recreation by rich people who had not the slightest relation to science. "Electric shows" were put on in the courts of European rulers. The Russian empress Catherine the Second took a great fancy to such things. Electric machines were built and electric sparks produced.

Be all this it may, but the science of electricity began from these simple experiments. And not only because the attraction of electrified bodies excited the imagination and was an impetus in the search to resolve the riddle, whereas, say, elastic forces are so common that they can hardly give rise to any kind of emotion. The main point is that here we come right up against a direct manifestation of one of the principal laws of interaction of charged bodies. This was easier to establish than it was to understand the interaction of atoms which make up electrically neutral bodies.

When at the beginning of the chapter we tried to follow through a series of questions and answers about the origin of elastic forces and then stopped at the beginning, we need not have stopped. We could have gone on to a discussion of atoms, their structure and the forces that act between them. Simply the exposition would not have been as convenient. But to suppose that we could arrive at the discovery of the basic laws of electromagnetic interaction through a study of the forces of elasticity is fairly

incredible. We might just as well suppose that man could have first invented the automobile and then, gradually simplifying it, come down to the cart, and finally the wheel. True, the city child finds it easier to understand a cart if we start explaining matters in terms of an automobile.

A person in the twentieth century, even if far removed from the scientific world, knows that it is easier to explain the motion of a stone than the motions of a cat. But to many people, even now, the forces of elasticity (why a football bounces, say) appear to be simple and clear-cut, whereas the attraction of pieces of paper by a comb or the attraction of two magnets is a pure riddle. Actually, it is just the reverse. The mysterious forces are the simpler, and the customary forces of elasticity are better grasped if reduced to the manifestation of "unusual" forces. That is exactly what we shall do a bit later.

Up to the middle of the 18th century, no great advances were made in the study of electricity. Two kinds of electricity were found to exist: positive and negative, and the possibility was discovered of transmitting and accumulating electricity; lightning was properly interpreted as a gigantic electric spark jumping between two clouds or between a cloud and the earth. And, finally, we came to the first practical application of these facts: Franklin invented the lightning rod. It was found that a pointed metal rod mounted on the top of buildings protected them from lightning strokes. The impression this made was tremendous. The lightning rod was all the rage. Women's hats were even fashioned with this modish rod (which brings to mind the fresh and startling hat models and hairdos after the first Soviet Sputnik went up). A curious sidelight is that King George III insisted that the lightning rods on his palace have rounded tips and not pointed ones, as suggested by that republican Franklin, who played a prominent role in the fight against England for the independence of the American colonies. This was too much for the President of the Royal Society, and he resigned.

Only after the great success enjoyed by Newton's mechanics was it possible to discover the exact law of interaction of stationary electrified bodies (or electri-

cally charged bodies, as the phrase usually goes). This law was first discovered not for separate elementary particles, the existence of which was not known at the time, but for large charged bodies. We now know that in electrification by friction, the charged particles of greatest mobility (electrons) move from one body over to the other. As a result, the body that loses the electrons becomes positively charged, while the one with excess electrons is negatively charged.

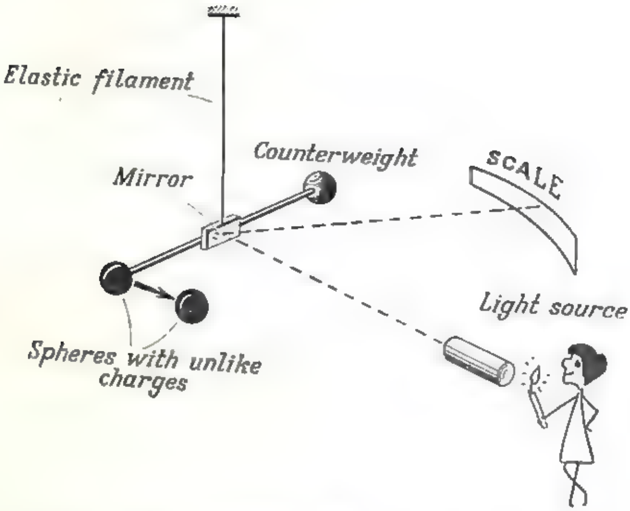
Problems of science

Newton's mechanics, the discovery of the forces of universal gravitation and the interpretation of planetary motion by that law produced such a tremendous impact on the minds of all scientists that workers in other branches of physics tried to discover laws patterned after Newton's. Scientific thought took the right tack. Instead of fruitless efforts to think up some kind of intangible mechanism generating forces that operate at a distance between charged bodies, attempts were made, experimentally, to find the quantitative form for a given type of interaction. One cannot overestimate the significance of this revolution in the approach to nature studies. Without doubt it was one of the greatest revolutions in natural

science. It began before Newton's time, as we pointed out at the beginning of the chapter on gravitational forces, and continued long after his death. The gist of this upheaval lay in the fact that the problem of science is not to reduce unusual "obscure" phenomena to ordinary, common sense phenomena. The problem of science is to find mathematically expressible general laws of nature that encompass a vast range of facts. Attempts were made to explain ordinary things on the basis of these laws, things so commonplace as would appear not to require any explanation. This was a direct challenge to "common sense". It was a challenge, which in such theories as relativity and quantum mechanics, even contradicted common sense. Unfortunately, this trend is not even today fully accepted by all. Time and again, naive questions are asked. It is no easy thing to overcome, for this is a revolution in thought comparable to that which occurred in the mentality of the wildman when he had to give up treating himself with such comprehensible methods as chasing out "evil spirits" and had to go over to such "mysterious" measures as observing the rules of cleanliness, boiling water, vaccination, antibiotics, etc. It appears that we have to chase out not human-like creatures, but microbes and viruses that cannot even be seen in a microscope.

Coulomb's law

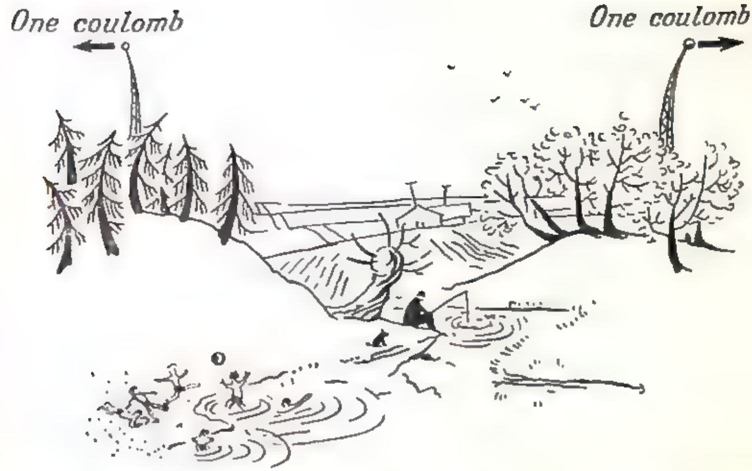
The discovery of the interaction of stationary (relative to one another) electric charges was made under the direct impact of Newtonian ideas, his law of universal gravitation for one. This discovery was made without great difficulty. In the middle of the 18th century, it was already suggested that the law of interaction of charges was analogous to that of gravitation. The first to prove it experimentally was the Englishman Cavendish. But this celebrated scientist was just as noted for his eccentricities. He was fanatically devoted to science. In order to save time he communicated with those about him in the household by a system of established signs. He did not publish his researches into electricity. For over a hundred years his manuscripts lay on the shelves of the Cambridge University library until Maxwell extract-



ed them and published them. By that time the law of interaction of charges had been discovered in France by Coulomb and to this day bears his name.

Coulomb arrived at his goal in a simpler yet less rigorous way than did Cavendish. Let us take a look at Coulomb's experiments.

The discovery of Coulomb's law was simplified by the fact that the forces of interaction between charges were large. Which meant that sensitive apparatus was not needed, as in the case of verifying the law of gravitation under terrestrial conditions. A simple device called the Coulomb torsion balance was enough to discover how stationary charged bodies reacted. The balance was simply a rod suspended by a slender elastic wire, at one end of which was a charged metal ball and at the other a counterweight. Another ball was fixed in place close to the balance. The interaction force was measured by the torsion of wire and a study was made of the dependence of the torque on the distance and magnitude of the charges. Force and distance were readily measured. The sole difficulty was with the charge. Coulomb's approach was simple and ingenious. He varied the magnitude of the charge of one of the balls by 2, 4 and so on times, connecting it with a similar uncharged ball. The charge was thus distributed equally between the balls, which reduced the magnitude of the charge under study in a definite



ratio. At the same time variations in the force were measured.

Coulomb's experiments led to the discovery of a law which was remarkably like that of gravitation: The force of interaction between stationary charged bodies is proportional to the product of their charges and to the inverse square of the distance between them, with the same reservation that Newton's law demands—Coulomb's law holds only for point charges, that is, charges whose geometric dimensions are small compared with the distance between them. Generally speaking, the force depends on the geometric dimensions and the shapes of the charged bodies. This force is usually called Coulomb's force.

The discovery of Coulomb's law made it possible for the first time to regard charge as a definite quantity to be measured.

This required a unit of measure. And Coulomb's law yields such a unit. True, it is impossible to make a standard charge (like a standard of length for the metre) due to leakage. The natural thing would be to take for the unit the electron charge (which has been done in atomic physics), but at that time nothing was known about the discrete structure of electricity. The unit was that charge which acts on an equivalent charge in vacuum at a distance of one centimetre with a force of one unit (the dyne). The practical unit is the coulomb, which is 3×10^9 times greater. In the former system of units,

the electron charge comes out to 4.8×10^{-10}, an extremely small quantity.

Coulomb forces slowly fall off with distance and belong to the group of long-range forces, like gravitation.

Here the similarity ends and the differences begin. First is the existence of charges of two different signs (gravitational mass is always positive). Electric charges exhibit both attraction and repulsion. Coulomb forces do not act between neutral bodies and are thus not so universal as are the forces of universal gravitation. Their universality is limited to the fact that the law holds both for macroscopic bodies and individual elementary particles. This was found to be the case immediately after such particles were discovered. From the present viewpoint, the Coulomb law is valid for macroscopic charges precisely because it holds for elementary particles.

Another very important peculiarity of Coulomb forces is their magnitude. The electric forces operating between individual elementary particles are immeasurably greater than the gravitational forces. Two charges of one Coulomb each would repulse with a force of 918 kilograms at a distance of one kilometre. However, the interaction between charged particles is so great that it is impossible to create a very large charge in a small body. The particles repulse each other with great force and cannot hold onto a body. And there are no other natural forces capable, under the given conditions, of balancing this Coulomb repulsion. That is one reason why perceptible attraction or repulsion of large charged bodies is not encountered in nature. What is more, charged bodies have a great tendency to neutralize. They avidly take up charges of opposite sign.

Most bodies in nature are electrically neutral. Incidentally, the earth itself has a negative charge of about 6×10^5 coulombs. In pure form, Coulomb forces operate mainly inside neutral atoms and in charged atomic nuclei. But about that later.

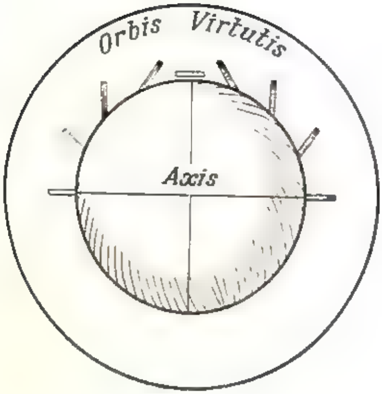
The important thing is that Coulomb's law was the first concrete step in the study of the properties of electric charge and thus in explaining the meaning of the very notion "electric charge". The presence of an electric charge in elementary particles or bodies means that they interact in accord with Coulomb's law.

4

Magnets interacting

Sometime in his life, everyone has found amazement and enjoyment with magnets. They act through space, through a void (since obviously the air is of no help!) and pick up pieces of iron, making whole strings of needles and tacks. No less remarkable is the behaviour of the magnetic needle of a compass that invariably points north no matter how we turn the compass to get it out of kilter. Possibly only the unusual ability of the top can compare with the magnet as food for the imagination.

The attraction of magnets resembles the action of electrified bodies at a distance. No wonder for centuries they were confused. It was Gilbert who demonstrated, at the end of the 16th century, that they are not the same. The magnet does not need to be rubbed to attract objects. And it continues to operate—unlike electrified bodies—if it is not heated considerably or violently shaken.

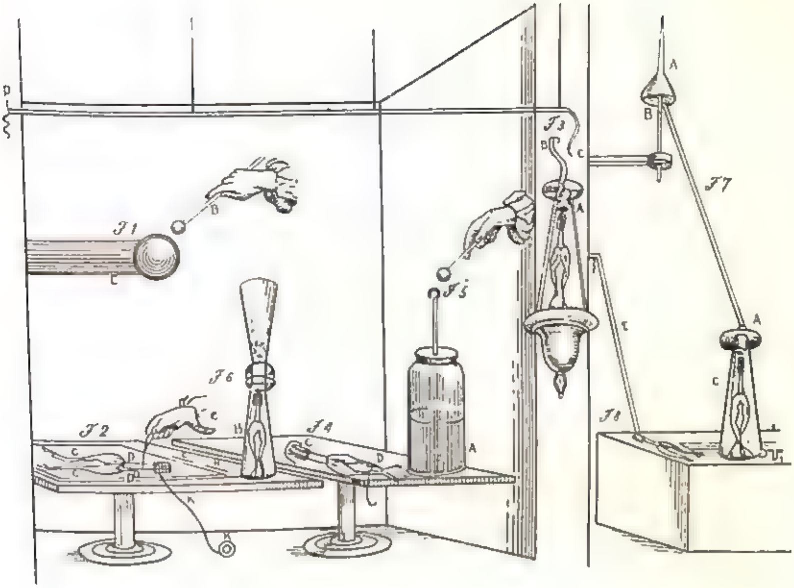


Magnets can attract and repulse like charges. But the strange thing is that no one is able to separate the north pole from the south and obtain an isolated magnetic pole, though great efforts have been made to do so.

The attraction of magnets is usually far greater than that of electrified bodies. That may be the reason why magical properties were attributed to them, while the weaker electrical attraction was not found good enough for such purposes. It was believed, for instance, that a magnet could cure diseases, reconcile husband and wife and do many other things.

As in the case of electrical attraction, for a long time there were no scientific inquires into the properties of magnetic interaction. Then again there was the opinion that a magnet would cease to function if rubbed over with garlic. Gilbert was the one who started studying magnets on a real scientific basis. It was Gilbert who conjectured that the earth was a gigantic magnet and for that reason made a magnetic needle orient in a specific manner. Gilbert was able to substantiate his conjecture experimentally by magnetizing a large iron sphere and observing its action on a needle. The position of small magnets with respect to the sphere was illustrated in Gilbert's book *Magnetism and Magnetic Bodies*.

The interaction of magnets was quantitatively studied by Coulomb, who used the same method as in the study of the interaction of charges. Coulomb found the law



of interaction of poles of long magnets, considering poles as points of concentration of magnetic charges, the analogues of electric charges. This law turned out exactly the same as the law of interaction of electric charges. Coulomb explained the impossibility of separating the north and south poles of a magnet as due to the fact that magnetic charges of opposite sign within molecules of matter cannot move freely from one molecule to another.

One might think (as Coulomb himself thought) that this was a fundamental law as in the case of the interaction of stationary electric charges. Coulomb introduced a new magnitude—the magnetic charge—and considered the newly discovered law of interaction of magnetic charges to have exhausted the problem of magnetism. And there was no apparent reason to doubt this, for Coulomb had acted according to pattern—Newton's law for the interaction of gravitating masses. Why indeed should this pattern work in one case and not in another; lead to the discovery of a new fundamental law in one sphere and not in another?

Electric current and "animal electricity"

Actually, everything was much more involved. Nature, in the fullness of its generosity, surprised investigators once again. Human imagination has a hard time catching up with mother nature. Magnetism was resolved from quite a different sphere. This occurred after we learned to generate electric current—the flow of electric charges—of considerable intensity and for appreciable lengths of time. The history of this discovery is of definite interest and is associated with the search for what was called "animal electricity".

Everything began with a discharge from a Leyden jar, the first capacitor. The man who discovered this phenomenon, Musschenbroek, was the first to test on himself the action of an electric discharge. Here is the way he put it: "The hand and the whole body is struck in such a terrible fashion that it is hard to describe. In a word, I thought the end had come." His advice to friends was "never repeat this new and terrible experiment".

Actually, however, this experiment is not so terrible. The brief electric current produced in a discharge of the jar is not dangerous to life. Be that as it may, the physiological action of an electrical discharge immediately attracted attention. There were many valuable observations and there was a spate of primitive theories that accounted for life and disease and death by electricity. Interesting and trustworthy discoveries were mixed with all manner of fanciful delusions. For example, a correct explanation was given for the effects of the torpedo fish and other electric fishes as being similar to the discharge of the Leyden jar. But along with this real "animal" electricity, people were discovering "electric" people, birds, and domestic animals. The confusion here was due to electricity being produced in the rubbing of clothing, fur or feathers.

In this atmosphere, some carefully thought out experiments of the outstanding experimenter Galvani yielded a fundamental discovery. True, Galvani himself was not able to interpret his own experiments correctly, but Volta who repeated them made a great discovery that straightway spurred the study of electromagnetism.

The first discovery was an accident. As Galvani wrote: "I cut and prepared a frog as indicated in Fig. 2, Table 1, and having in view something quite different, I placed it on the table where an electric machine, Fig. 1, was located, disconnected completely from the conductor of the latter and at quite a large distance from it. When one of my helpers accidentally touched the femoral nerves of this frog with the sharp end of a scalpel, all the muscles of the limbs immediately began to contract violently as if caught by strong tonic cramps (This occurred when a spark was extracted from the conductor of the machine—*Authors' remark*). It was then that I was gripped by an incredible urge and passionate desire to investigate this phenomenon and extract what was hidden there."

Galvani soon noticed that a frog's leg connected with a lightning rod contracted every time lightning struck and even when thunder-storm clouds were approaching.

What was observed in these experiments for the first time was electromagnetic induction, later discovered by Faraday. At that time it was still too early to interpret it correctly. The discovery that spurred the development of electromagnetism was something quite different.

Galvani attempted to detect the action of atmospheric electricity in fair weather. For this purpose he hung a prepared frog on an iron fence with a copper hook passing through the spinal chord. Pressing the hook to the iron slats, Galvani observed violent contraction of the muscles. Fortunately, he realized that this had nothing to do with atmospheric electricity. Contraction was observed at all times when two dissimilar metals in contact were pressed to the frog's leg.

Knowing that muscle contraction occurs in the case of an electrical discharge, Galvani decided that he had discovered animal electricity generated in the organism. Galvani reasoned that the metal conductor enables the electricity to move rapidly from certain parts of the muscle to others, thus causing contraction.

The correct explanation was given by another Italian, Volta, who, on this basis, produced the first source of direct current. Therein, essentially, lay the significance of Galvani's discovery for physics. Volta brilliantly

conjectured that the frog's legs were simply a sensitive "animal electrometer", more sensitive than any other and nothing more. The source of electric current was the two dissimilar metals brought into contact with the electrically conducting liquid of animal tissues. This is where Volta got the idea of the first galvanic cell: a set of copper and zinc circular plates with an interposed cloth soaked in salt water. This was the "voltaic pile", which, as Arago wrote, was "the most remarkable instrument ever invented by man, not excepting the telescope and the steam engine".

The curious thing is that neither Volta nor any of his contemporaries had the slightest idea of how and why this device worked. Which, incidentally, was not so very important for the development of science at that time. The main thing was that the voltaic pile generated direct electric current, that is, it could make electric charges in a conductor move. The explanation of this action was a long time in coming, and we shall not go into it here.

Oersted's discovery

The voltaic pile was a virtual cornucopia. New discoveries came one after the other. Davy used current to break down an alkali and obtain metallic sodium and potassium. Petrov discovered the electric arc. Finally, Oersted, in 1820, made a very important discovery. Placing a compass needle close to a current-carrying wire, he found that it turned.

Now this was no accidental discovery. As early as 1807 Oersted had made up his mind to find out whether electricity had any effect on a magnet or not. Faraday wrote that the persistence with which he pursued his aim was rewarded by the discovery of a fact whose existence was not even surmissed by anyone except him, but which, when made known, immediately attracted the attention of all who could appraise its significance and value.

A direct link-up was finally made between the fortuitously discovered remarkable ability of pieces of iron to be attracted at a distance (made by shepherds of the ancient world) and the jerking of a frog's leg in Galva-

ni's experiments. Magnetism and electricity were related in a profound way, and this was demonstrated by experiment. It may be added that the compass needle remained totally indifferent to charges at rest. Only moving charges were capable of generating "kindred emotions". Magnetism is associated not with static electricity but with electric current.

Magnetic interaction is the interaction of electric currents

Oersted's discovery almost immediately permitted solving the riddle of magnetism and at the same time made for the discovery of yet another fundamental type of interaction of electric charges in addition to the Coulomb action. All this was done by one man, Ampere, during just a few months after he had learned of Oersted's experiment. The chain of reasoning of this brilliant mind recorded in his communications to the French Academy of Sciences that followed one after the other is extremely interesting. First, under the impression of observations of a compass needle turning close to a current-carrying conductor, Ampere decided that the magnetism of the earth was due to currents flowing round the earth from west to east. The main step was taken. *The magnetic properties of bodies could be accounted for by currents circulating in them.* Ampere then came to the general conclusion: the magnetic properties of any body are determined by closed electric currents circulating within it. This decisive advance from the *possibility* of explaining the magnetic properties as due to currents to the categorical *assertion* that magnetic interaction is the interaction of currents is an indication of the scientific audacity of Ampere.

On Ampere's hypothesis, elementary electric currents circulate inside the molecules that make up bodies. If these currents are arranged at random, their action cancels out and no magnetic properties are exhibited by the body. In a magnetized state, the elementary currents in the bodies orientate in a very definite manner so that their effects combine.

Where Coulomb had seen indivisible magnetic poles of molecules, there were found to be only closed electric currents. The indivisibility of magnetic poles lost its mysteriousness completely. There are no magnetic charges and therefore nothing to divide. *Magnetic interaction is not caused by specific magnetic charges, as in the case of electric interaction, but by the motion of electric charges—by current.*

The fruitful unitary concept of the forces of nature was perhaps never exhibited more distinctly than in the formulation of the basic laws of electromagnetism. Inspired by this idea, Oersted brought a compass needle up to a current-carrying conductor, and Ampere mentally perceived electric currents within the magnetic piece of iron. This same idea later led Faraday to the great discovery of electromagnetic induction.

Ampere's law

Ampere not only conjectured that studies of magnetic interaction should begin with an investigation of the interaction of electric currents, but set to work immediately with experiments to prove it. One thing he found was that currents flowing in one direction attract, while those flowing in opposite directions repulse each other. Mutually perpendicular conductors have no effect on one another.

Ampere's intense efforts were finally crowned with success. He discovered the law of mechanical action between electric currents, thus resolving the problem of magnetic interaction. The law of interaction of magnetic poles, which Coulomb had considered fundamental, was really only one of a large number of corollaries of Ampere's discovery. About Ampere, Maxwell wrote that both theory and experiment came mature and fully armed from the head of this "Newton of electricity". These investigations, said Maxwell, were complete in form, ideally precise and summarized in a formula from which all phenomena are derivable and which for all time should remain the fundamental formula of electrodynamics.

We shall not go into detailed descriptions of the experiments that led Ampere to the discovery of current interaction, as we did for the much simpler case of the interaction of stationary charges. And there is no need to state Ampere's law for currents as he himself did. Now, since electric current is nothing other than a flow of electric charges, current interaction reduces to the interaction of moving charges. Thus, in addition to Coulomb interaction, which is determined solely by the magnitude of the charges and the distance between them, a new type of interaction comes into play when the charges are in motion. It is determined by the charges and the distance *and* the rate of motion of the charges. *This was the first discovery of fundamental forces dependent on velocity!*

*The force of interaction of moving charges is proportional to the product of these charges and the inverse square of the distance between them, as in Coulomb's law; but, in addition, it is also dependent on the velocities of the charges and their direction of motion.** It is in the discovery of this law that the meaning of all previous efforts lies.

Magnetic forces are radically different from electric forces in yet another respect. They are not of a central nature, like Coulomb and gravitational forces. This was evident even from Oersted's experiments, for the compass needle was not attracted to the wire or repulsed from it, but only turned. The force discovered by Ampere acts on moving particles in a direction perpendicular to their velocities.

The forces of magnetic interaction of particles are much weaker than Coulomb forces under ordinary conditions. Only at particle velocities approaching the speed of light do they become comparable. Nevertheless, the forces of current interaction may attain a great magnitude. Recall that these are the forces which rotate the armatures of enormous electric motors. The more powerful Coulomb forces are, on the whole, technically useless. The whole point is that we can create very large currents

---

* These forces supplement the Coulomb forces, which do not vanish in the case of motion.

that is, we can put vast quantities of electrons in conductors into motion (true, relatively slow motion). But it has not been possible to generate very large electrostatic charges. Strange as it may seem, magnetic interactions are important only technically (recall electric motors), whereas in nature, as we shall see later on, they play a very modest part as compared with Coulomb forces. This is because these are forces of current interaction, which in nature rarely attain a considerable magnitude.

Ampere's discovery extended our conceptions of electric charge. Charge is found to possess a new and fundamental property: the ability to interact with forces that depend on the velocity of motion.

5

Close-range action

The laws of interaction of stationary and moving charges were found. But it was still not clear how force is conveyed from one charge to another, just as the discovery of the law of universal gravitation did not explain the nature of the forces of gravitation. We have already spoken of the problems that are common to both gravitation and electromagnetism. However, these problems are so important that we had better examine them again and in more detail. All the more so since they came to the fore for the first time (historically) during studies of electromagnetic phenomena.

Perhaps no one has made the point so clear as Maxwell in a paper dealing with action at a distance.

Maxwell says that if we observe the action of one body on another a certain distance away, then before assuming

that this action is direct we are inclined first to investigate and find out whether or not there is some material connection between the bodies in the form of fibres, rods, etc. If such connections exist, we prefer to explain the action of one body on another by means of these intermediate links.

This calls to mind the old autobus where the driver opened doors via a series of interconnected links. In modern buses, compressed air is used: it moves through pipelines and acts on a cylinder that controls the door mechanism. The same purpose may be served by using an electromagnet and sending signals to it over wires. In all these three modes of door-opening there is one thing in common: there is a continuous connecting line between the driver and the door, and some specific physical process occurs at each point in the line. Action is transferred via this process and with a finite velocity, not instantaneously.

Thus, Maxwell concluded, action between bodies at a distance may be accounted for in many cases by the presence of certain obvious intermediate agencies that convey the action. Would it not, then, be reasonable to presume the existence of some intermediate agency in cases where there is no apparent medium, no intermediary between interacting bodies? *Therein lies the essence of the conception of close-range action.* Otherwise we

would be forced to say that a body acts in regions where it does not exist.

For one who knows nothing of the properties of air, a ringing bell would seem to operate directly on our ears, and the transmission of sound by an invisible medium would appear to be quite incomprehensible. However, in this case we can follow every detail of the propagation of sound waves and compute their velocity.

Maxwell goes on to say that many thinkers began to imagine invisible emissions surrounding the planets (recall our talk about gravitational forces) and magnets, and invisible atmospheres about electrified bodies. These speculations were often ingenious but their one great fault was that they were fruitless and offered nothing tangible to science.

Action at a distance

This continued until Newton established his law of universal gravitation, for which, unfortunately, no explanation was given. Subsequent studies of the solar system so gripped the imagination of investigators that for the most part they inclined to the view that a search for some kind of mechanism was useless. The idea arose of direct action at a distance through empty space. Bodies were thought capable of "feeling" the presence of one another without any intervening medium. Attempts were frequently made to support the action-at-a-distance notion by Newton's authority, but as we have already pointed out, this was not true.

The adherents of action at a distance were not disturbed by the idea of action in the absence of any bodies. They pointed to a magnet and said: Isn't it obvious that a magnet attracts bodies across empty space and that the force of attraction is not even perceptibly diminished if the magnet is wrapped in paper or put into a wooden box. What is more, even if we believe a certain action to be due to direct contact, this is actually not so, for even in the case of the most intimate contact between bodies, a small interval remains. Isn't it clear that a weight

suspended by a thread does not tear it even though there is empty space between the atoms that make up the thread bodies. Action at a distance is, therefore, not only not impossible—it is the only universal mode of action. Close-range action exists only in the minds of its adherents, and not in nature. This is because the view is based on the crude experiments of pre-scientific times when contact was considered necessary for interaction, and when people did not understand that no direct contact ever occurs, there being action at such small distances that the existing imperfect methods of observation could not record it.

Quite strong reasoning, as you see. And strengthened too by the marvellous successes achieved by such supporters of action at a distance as Coulomb and Ampere.

If the development of science took a direct path, there would be no doubt whatsoever about the ultimate victory of action at a distance. Actually, however, the advance of science is more in the way of a spiral than a straight line. Each fresh turn brings us back to roughly the same concepts, only at a higher level. That is what happened in the case of close-range action.

The progress made in the discovery of the laws of interaction of electric charges was not integrally linked up with the concept of action at a distance. This is because experimental investigation of the forces themselves does not at all necessarily presuppose definite conceptions about how the forces are transmitted. The first step was to find a mathematical expression for the forces. Explanations could come later. The success of the action-at-a-distance adherents made it patently clear that it was useless to attempt to explain the fundamental laws of nature through the use of pictorial mechanical models borrowed from the gross world of everyday life.

Faraday's electromagnetic field

A decisive step in the realm of close-range action was taken by Faraday—creator of the basic ideas of electromagnetic theory—and the final step was made by Maxwell. *According to Faraday, electric charges do not act*

on one another directly. Each one of them (if it is in motion) sets up an electric field and a magnetic field in the surrounding space. The fields of one charge act on the other and vice versa.

Lying at the heart of Faraday's views on the electric field was the notion of lines of force that extend out in all directions from electrified bodies. These lines, which indicate the direction of action of an electric force at each point, were not new. They had been observed and studied as an oddity.

If the elongated crystals of a dielectric (say, quinine) are thoroughly mixed in a viscous liquid (castor oil, for example), the crystals will align themselves in chains near charged bodies, forming all kinds of patterns that depend on the distribution of charge.

It is possible to trace force lines near the surface of the earth just before a thunderstorm.

In exactly the same way, we can observe magnetic lines of force near a current-carrying conductor by means of iron filings.

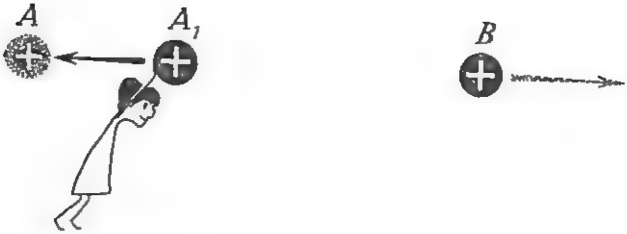
Faraday was the first to reject force lines simply as a method of embracing at a glance the resultant direction of action-at-a-distance forces from electrified bodies or currents at various sites: the compound resultant of simple laws. On Faraday's view, lines of force afford a real picture of actual processes occurring in space near electrified bodies or magnets. He imparted to the concept of lines of force a peculiar clarity and exactitude. Faraday's distribution of force lines yields a picture of an electric field near charges or of a magnetic field near magnets and conductors.

Maxwell said that Faraday visualized all space as permeated with lines of force. Faraday saw a mediating agency wherever mathematicians imagined centres of stress of long-range forces. And where the latter saw nothing except distance, being satisfied that they had discovered the law of distribution of forces acting on electric fluids, Faraday sought the essence of actual phenomena occurring in a medium. Though Faraday was not mathematically minded and could not follow the reasoning of such brilliant mathematicians as Ampere, yet with the aid of lines of force he was able to grasp the most involved problems of electrodynamics. There can be no doubt that it was precisely these ideas that led him to a number of exceedingly important discoveries.

Contemporaries, excited by the success of the studies of Ampere and other authorities on action at a distance, were rather cool about Faraday's ideas, though they did follow his experimental discoveries with interest. As one of them wrote: "I can't understand how anyone seeing the coincidence between experiment and calculations based on the law of action at a distance would hesitate for a moment in his choice between this clear-cut comprehensible action and a hazy, nebulous thing like lines of force."

Electromagnetic fields exist

However, the adherents of action at a distance could not for long be proud of the mathematical elegance and exactitude of their theories. The great Maxwell was able



to express Faraday's ideas in exact quantitative form, which is so necessary to physics. He wrote his famous system of equations of an electromagnetic field. It was found that the laws discovered by Coulomb and Ampere find their most complete, profound and mathematically elegant form precisely in the language of fields. Since that time, the concept of an electromagnetic field began to gain in popularity. Complete victory came later, roughly fifty years after the formulation of Faraday's basic ideas.

Maxwell was able theoretically to prove that electromagnetic interactions are propagated with a finite velocity, the velocity of light in empty space: $c = 300{,}000$ km/sec. Which means that if we move some charge A slightly, the Coulomb force acting on charge B will not change *instantaneously, but only after a lapse of time* $t = \frac{AB}{c}$. This was a fundamental result that utterly wiped out the notion of action at a distance. There actually is a process (which we will look into later on) between charges in a vacuum, as a result of which interaction between the charges is propagated with a finite velocity. True, an experiment of this nature is difficult to perform due to the extreme velocity of propagation of the process. But there was no need for it, since from Maxwell's theory there followed a fundamental fact: an electromagnetic field possesses a peculiar inertia. Sudden changes in the velocity of a charge make the accompanying field detach itself from the charge, just like people standing in an accelerating train are thrown out of balance if they are not holding on to something. The fields that become detached from the charge begin an independent existence in the form of electromagnetic

waves. This is a very familiar thing, since it underlies the operation of every radio station, the purpose of which is to broadcast electromagnetic waves. And if the station suddenly stops operating, the emitted electromagnetic waves will continue to wander in space for a long time until they are absorbed by bodies. In this case, the electromagnetic field is something just as real as, say, the table you are sitting at, and we can no longer brush aside a field as something that complicates and confuses simple things, as the action-at-a-distance people attempted to do.

The idea that a body could not act directly at a spot where it was absent (which idea, when it first appeared, seemed to be a piece of self-contradictory nonsense) was disproved experimentally despite the fact that the very development of science had at one time appeared to demand its recognition and that the thought-curbing dogmas of short-range action should be rejected.

6

An agonizing question

Just what is an electric field and a magnetic field? This is always the thing that worries a person striving to understand the basic quantities that modern physics deals with, a person who has either lacked time to make a detailed study of the matter or who has lost all hope of working in that field. This is the most frequently asked question in the mail of all popular-science magazines and publishing houses. The electric charge arouses somewhat less interest, yet it is just as complex as the field. This is probably because a charge is associated with something tangible—an electrified body—while a field is not.

The questions come from different people with a school-level knowledge of fields. They are puzzled that no satis-

factory definition is ever given. College students never ask such questions, though. Apparently, they are aware that a few sentences will not suffice, or they know that explanations can be got elsewhere.

The electromagnetic field and the ether

The situation is indeed involved. The first notions about force lines suggested by Faraday and followed up by Maxwell were born in a period when Newtonian mechanics was triumphant. It was believed to be universal and all-embracing. Newton's postulates had long since ceased to be regarded as hypotheses based on an experimental foundation. They had become almost self-evident.

Neither Coulomb nor Ampere had ever thought of departing from the Newtonian stand. They were only studying new types of force. And Newton's theory allows for any kind of force.

Essentially, Faraday took the same stand, with the important difference, however, that he did not accept action at a distance. Faraday was not satisfied with writing formulas that permitted expressing electromagnetic forces in terms of distance, velocity and so forth. He attempted to get a picture of the actual *mechanism* of the origin of these forces. Note that mechanism is meant in a very literal sense. This (together with experiments with iron filings and pieces of dielectric) led Faraday to the concept of lines of force very reminiscent of ordinary elastic fibres (even though invisible and beyond direct perception by the sense organs).

Yes, no matter how paradoxical it may seem today, both Faraday and Maxwell stood for a *mechanical* explanation of electromagnetic phenomena!

Starting with the hypothesis that space is filled with an all-permeating medium called the ether, they attempted to reduce all electromagnetic phenomena to mechanical motions in the ether, to mechanical stresses within it. Much in today's theory leans in that direction. People still write (though the words have a different content) about tensions associated with an electromagnetic field, about fluxes and vortices.

Scientific discoveries often take strange turns. Fourier, for instance, created a correct mathematical theory of thermal conductivity, starting from the utterly erroneous notion of caloric, the hypothesized fluid carrier of heat. Yet we still use this theory. Faraday and Maxwell built a harmonious theory of electromagnetism on the basis of mechanical views.

In this latter instance, the logic of development of the ideas was amazing. The ether was found to be nonexistent. One could reconcile himself to the exotic properties ascribed to it: fantastic elasticity and negligible density and viscosity. But there gradually came to light circumstances which not only became a challenge to the demands of pictorialness (in itself acceptable), but to the logical integrity of the theory as such. For example, in some experiments the ether (if it existed) was to be entrained by moving bodies, completely entrained. From other experiments it followed that the entrainment could only be partial. Finally, there were experiments in which there was not the slightest entrainment! This hypothetical medium turned out to be quite elusive.

The ether and the theory of relativity

These contradictions gave a jolt to the customary and hardening concepts of physicists as regards the ether. The mechanical ether was finally buried by Einstein's theory of relativity. *It transpired that not only was it totally impossible to construct a satisfactory mechanics of the ether, but even to detect motion with respect to it was out of the question.*

The remarkable thing is that this did not disturb a single element in Maxwell's elegant mathematical formulations of the laws of an electromagnetic field. The equations remained intact! It might be better to say that they retained their original form, but their meaning became different, that is, the meaning of the terms "electric field," and "magnetic field". Thus, lines of force in today's theory give a representation of the distribution of a field in space, but in no way do they signify taut "strings" of an invisible mechanism. In this sense, they

are no more real than the meridians or parallels of latitude on a globe.

We may say that the adherents of action at a distance were right in one respect after all. They were wrong in rejecting an intermediate agency that gives rise to interaction. But they were right in ridiculing attempts to account for these interactions by some kind of imperceptible mechanism that should be so constructed as to yield only what we observe in reality and nothing more. If there is no ether, then we cannot hope to reduce electromagnetic phenomena to mechanics, an involved and very peculiar type of mechanics but still Newtonian mechanics.

If that is the case, then while studying electromagnetic fields we encounter some kind of matter (and there is no reason to doubt that an electromagnetic field is material) that *does not obey Newton's laws*! This matter is described by its own specific laws given mathematically by Maxwell's equations.

This was a fundamental discovery. A turning point in the history of science: different kinds of matter exist and each type is described by its own laws which are totally distinct yet with certain points in common. This idea was born the day that the mechanical ether was discarded.

What does explain mean?

But what kind of matter is this? How do we now define electric and magnetic fields? Our job now is to go into a rather lengthy and boring line of reasoning with the sole purpose of justifying ourselves for not being able to give a satisfactory definition of field.

Recall various ways one would try to explain what a certain thing is. One way, of course, is to point it out, and then with one's sense organs it is possible to find out all kinds of things. If the object is not at hand, or if it is invisible (this is a second way), one can give a detailed account of its properties. Finally, if the need arises, you can describe its structure, what it consists of.

We are accustomed to all these methods, but one is preferred in one instance, another in a different situation.

For example, to explain what a giraffe is might be rather difficult, but one glance would obviously be sufficient so as not to confuse it ever again. Now it is quite different to learn about the feelings of a person lost in an avalanche of snow in the mountains. Here a story would be better. For an atom, the simplest way is to learn about its structure. The first of our methods couldn't be applied anyway, the atom being too small to see or touch.

Frequently all three methods are equally possible and one can select any according to the nature and degree of interest in the matter at hand. For instance, if you wish to learn about glucose, you can take the encyclopedia or some other manual and read a description of its properties. In this way you will find that glucose is made up of colourless crystals that melt at 146°C, that it is roughly half as sweet as beet sugar, and so on. The list of properties is quite extensive.

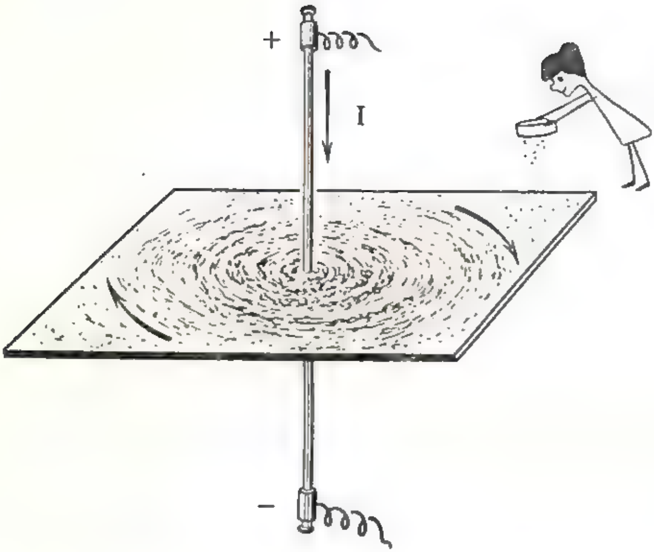
Then again you can take up its structure and find that glucose consists of carbon, hydrogen and oxygen. One molecule of glucose contains six carbon atoms, six oxygen atoms and twelve hydrogen atoms linked together in a specific fashion.

Finally, one can simply get hold of a glucose crystal and see what it looks like.

A scientifically minded person will get the best and most satisfying conception by learning about the structure of an object. Particularly, if the structure gives a clue to its various properties. This essentially is the purpose of science.

Primary entities

However, there are entities that cannot be explained either by the first or the third method.These entities are beyond our sense organs and we can say nothing about their structure. Such are electric and magnetic fields. It is not so terrible that these fields do not act on our sense organs, though it is not so easy to be convinced of their existence if we cannot sense them directly. Atoms cannot be perceived by our sence organs, yet we get accustomed to them rather easily. The field is



quite a different and more complex matter in the sense that nothing can be said about its structure. This is an unusual situation indeed. It occurs only in the case of the most elemental of entities (the elements of the ancients) known today. There is nothing more primal than an electromagnetic field. And that is why we are unable to say anything about its structure.

At every stage in the development of science we come up against such elemental entities that cannot be broken down into component parts for the simple reason that such parts are unknown. Ancient philosophers distinguished four elements: water, air, fire and earth. Later they became atoms, and now elementary particles and fields. The only possible question is: Will simpler entities be discovered in the future that may be regarded as the component parts of fields and particles? So far nothing definite can be said.

A first warning: do not try to visualize a field as something simple. An elementary particle is easily pictured as a ball or something clear-cut and discrete in space. With a field we are inclined to associate something that permeates all space, like a liquid filling a vessel. Such

views were common at the close of last century. An electron was pictured as a charged sphere, while an electromagnetic field was visualized as stresses in a hypothetical medium, the ether. In actuality, these pictures are far removed from reality. On present views, an electromagnetic field exhibits properties typical of particles, while elementary particles in turn display typically wave properties. But we shall not go into details yet.

Basic properties of an electromagnetic field

Now we can discuss an electric (electrostatic, to be precise) field. All our ideas about electric fields are obtained from experimental studies of their properties. There is no other way of discovering such properties. *The principal property of an electric field is its ability to act on electric charges (both stationary and moving) with a certain force.* Judging from this action on a charge, we establish the existence of a field, its distribution in space, and we study its characteristic features.

An electric field is set up by electric charges. By convention, the lines of force of this field begin on positive charges and end on negative charges. Charges are the sources of a field. From the action of the field on a charge we can detect the field, and by studying this action we can introduce a strictly definite quantity that permits *measuring* the field. This quantity—the field intensity—is a force which acts on a unit positive charge.

The principal property of a magnetic field is the ability to act on moving electric charges with a definite force. Likewise, a magnetic field is created only by moving electric charges. The lines of force of a magnetic field encompass the currents in the form of closed lines without beginning or end.

Starting from the discoveries of Coulomb and Ampere, Maxwell formulated precise laws that determine the magnitude of electric and magnetic fields as a function of the spatial distribution of charges and currents.

The reaction of scientists to fundamental concepts

It is well worth adding a few words about the attitude of physicists themselves to such fundamental notions as field. To many, this definition of a field and the enumeration of its properties will appear to be rather deficient. Perhaps we should first of all make every effort to achieve clarity in the question of the field and attempt to obtain a detailed picture of its nature.

Scientists think otherwise. Of prime importance to scientists, as far as the information that we have about fields is concerned, are the tremendous opportunities we thus have for explaining an immense range of experimental facts. Of course this information includes precisely formulated mathematical laws that determine the configuration of a field depending on the arrangement of charges and their velocities, and not only the qualitative ideas that we have just given. Scientists are clearly aware of the fact that the situation here is the very same as in Newtonian mechanics. As you recall, mechanics does not care about the nature of the force, but wants to know what it is equal to and under what conditions it appears. And so too in the theory of the electromagnetic field, it is above all important to know how the field is acting on a charge and under what circumstances the field originates, not what the field in itself is. The only difference is that when we go beyond the limits of mechanics we can investigate the nature of forces, but we cannot do this with a field, at least for the present.

Studies of the nature of various mechanical forces consist essentially in reducing them to certain fields. The fields cannot, as we have already said, be reduced to anything more elemental, at least at the present time.

Sooner or later we will have more information about fields, but present information does not permit conjecturing about the mechanism of action of a field on charges. We shall have to be satisfied with this, for it is impossible to return to the early attempts at a mechanical interpretation of a field. Any attempt to grasp at once the "very essence" of a field instead of entering on the hard

and arduous and exceedingly important path of explaining specific phenomena on the basis of known properties and searching for new field properties might appear to be laudatory, it is really something we must rid ourselves of.

Today, we can't even say for sure that *in the future* scientists will discover more fundamental entities than fields and elementary particles.

Nature is inexhaustible in its properties. As Lenin pointed out, the electron, too, is inexhaustible. So also is the electromagnetic field inexhaustible in its properties. Hence the process of an ever increasing cognition of the properties of fields will never cease. But are today's most elementary structures divisible *ad infinitum*? The present assemblage of facts suggests a NO. If that is the case, then further progress in the study of fields and elementary particles will be associated solely with discovery of more and more properties. So far we have dealt only with the most important properties and have said nothing about what modern science knows concerning fields.

Let us now examine some more fundamental properties of an electromagnetic field.

7

New properties of an electromagnetic field

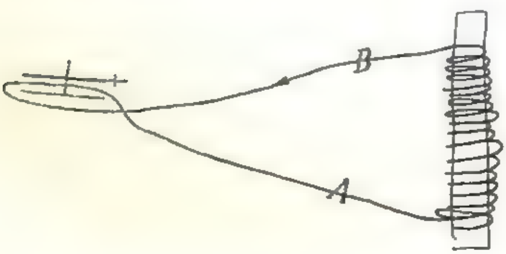
If stationary charges set up electric fields, and moving charges gave rise to magnetic fields, and nothing more, then the family of electromagnetic forces would be a much smaller one than we find it. What is more, it would not be possible to state with the same assurance that fields are just as much a reality as the existence of the authors of this book for the readers. Scientists of the close-range

orientation would be in the position of beginning writers who are not sure that they will find readers for their books.

The situation was changed radically by the discovery of fresh properties of electromagnetic interactions, which could not, without stretching a point, be interpreted in the language of action at a distance (as was done for the laws of Coulomb and Ampere). Electric and magnetic fields turned out to be intimately connected. Under certain conditions, a magnetic field is capable of generating an electric field without the aid of charges, and an electric field can give direct birth to a magnetic field. That is precisely the situation: a magnetic field generates only an electric field, while an electric field can produce only a magnetic field, which, true, can in turn generate an electric field. Something along such lines is seen in the world of insects when a caterpillar turns into a butterfly, and only a butterfly, and the latter produces only eggs, which yield caterpillars, which never have caterpillars as offspring directly, and the same goes for butterflies.

Electromagnetic induction

It was not by accident that the first and most important step in the discovery of this new aspect of electromagnetic interactions was made by the founder of the electromagnetic field concept, Michael Faraday, one of the world's greatest scientists. Faraday was absolutely convinced of the unity of electric and magnetic phenomena. Soon after Oersted's discovery he wrote in his diary (1821) about turning magnetism into electricity. Since then Faraday never ceased thinking about this problem. They say that he always carried a magnet with him in his



vest pocket to remind him continuously about the problem. And ten years later, as a result of persistent work and faith, the problem was solved. He made a discovery that lies at the basis of the generators of all electric power stations in the world that convert mechanical energy into the energy of electric current. Other sources, such as galvanic cells, storage batteries, thermocouples and photoelectric cells generate only a minute portion of the electricity we use.

Faraday reasoned this way: electric current is capable of magnetizing a piece of iron. All you need to do is put it into a coil. Isn't a magnet capable of producing an electric current or of changing the magnitude of a current? For a long time nothing was detected.

The following curious fact is indicative of the kind of incidents that often interfere in a discovery. Almost at this same time, a Swiss physicist Kolladon was also trying to obtain an electric current by means of a magnet. He used a galvanometer, the light magnetic needle of which was positioned inside the coil of the instrument. So that the magnet should not exert any direct effect on the needle, the ends of the coil into which the magnet was inserted in the hope of obtaining current were carried out into an adjacent room and there connected to a galvanometer. When he moved the magnet into the coil, Kolladon went into this room and was disappointed to find that the galvanometer indicated zero. If he had only observed the galvanometer and asked someone else to handle the magnet—the remarkable discovery would have been made. But he was alone, and the magnet at rest relative to the coil could remain there a hundred years without causing any current to flow in the coil.

Faraday encountered the same kind of difficulties because he attempted time and again to obtain an electric current by means of a magnet and with the help of current in a single conductor, but failed.

The discovery of electromagnetic induction, as Faraday himself called it, was made in 1831. Here is a brief description of the first experiment:

"Two hundred and three feet of copper wire in one length were coiled round a large block of wood; other two hundred and three feet of similar wire were interposed as a spiral between the turns of the first coil, and metallic

contact everywhere prevented by twine. One of these helices was connected with a galvanometer, and the other with a battery of one hundred pairs of plates four inches square, with double coppers, and well charged. When the contact was made, there was a sudden and very slight effect at the galvanometer, and there was also a similar slight effect when the contact with the battery was broken. But whilst the voltaic current was continuing to pass through the one helix, no galvanometrical appearances nor any effect like induction upon the other helix could be perceived, although the active power of the battery was proved to be great, by its heating the whole of its own helix, and by the brilliancy of the discharge when made through charcoal."

Thus, the first discovery was that of the induction of currents stationary relative to one another. Then, clearly aware that closing and opening correspond to bringing current-carrying conductors closer together and increasing the distance between them, Faraday experimentally demonstrated that current appears when coils are moved relative to one another.

Faraday was acquainted with the works of Ampere and understood that current is a magnet, and a magnet, in turn, is an assemblage of currents. On October 17, as he wrote in his diary, he detected an induction current in a coil when a magnet was being inserted into the coil or withdrawn from it.

In the course of one month Faraday experimentally discovered all the essential peculiarities of the phenomenon. Tyndall, a friend of Faraday, wrote that his mighty mind covered a broad field and hardly left a crumb of facts behind for other investigators. The final step was to express the law in a rigorous quantitative form and fully determine the physical nature of the phenomenon. Faraday himself detected the underlying factor that gave rise to an induction current in these outwardly different experiments. A current arises in a circuit when there is a change in the number of lines of force of the magnetic field that link the area bounded by this circuit (i.e., a change in the magnitude of the magnetic field linking the circuit). And the faster the variation in this number, the greater the current. How the number of lines of force is changed is totally immaterial. It may be due to a change in current

intensity (and hence in the field), or a moving together of coils, or the motion of a magnet.

Faraday not only discovered the phenomenon, but was first to demonstrate a model (rather imperfect as yet) of an electric-current generator converting the mechanical energy of rotation into current. This was a massive copper disk rotating between the poles of a strong magnet. Faraday connected the axis and edge of the disk to a galvanometer and saw that the needle was deflected. True, the current was weak, but the principle which was established enabled powerful generators to be constructed. Without them, electricity would still be a rare luxury.

The direction of an induction current and conservation of energy

An induced current immediately begins to interact with the current or magnet that generated it. If the magnet (or current-carrying coil) is brought near a closed circuit, the current induced in it will always repulse the magnet. Work has to be done in order to move it closer. When the magnet is taken away, attraction sets in. This rule was found by Lenz and holds in every case. Imagine a contrary situation: you push the magnet towards the coil and it goes right on into the coil and—we have a violation of the law of conservation of energy, because the mechanical energy of the magnet would be increased and at the same time there would appear a current, which in itself requires the expenditure of energy, since a current is able to perform work. Nature wisely arranged the direction of induced current so that the total energy would not change. The current induced in the armature of an electric generator interacts with the magnetic field of the stator and slows down the rotation of the armature. That is the only reason why the bigger the current, the more work has to be done to rotate the armature. Current is induced at the expense of this work.

It is curious to note that if the magnetic field of this planet of ours were very large and highly homogenous, fast moving conducting bodies on the surface and in the atmosphere would be impossible due to intensive inter-

action between this field and the current induced in the body. Bodies would move in a dense viscous medium and would heat up tremendously. Aircraft or rockets would never fly. A human being would not be able to move his arms or legs because the human body is a fair conductor.

If a coil in which a current is being induced is stationary relative to a proximate coil with an alternating current (say, a transformer), then in this case too the direction of the induced current is determined by the law of conservation of energy. This current is always in a direction such that the magnetic field it generates strives to reduce the change in current in the primary winding.

The nature of electromagnetic induction

As soon as Faraday discovered the law of electromagnetic induction, scientists attempted to put it in a rigorously quantitative form. Today it is hard to imagine the tortuous efforts required to state the law in terms of action at a distance. At long last, Neuman and Weber produced some extraordinarily complicated formulas which were rather nebulous as to physical content yet capable of describing experimental facts quantitatively. Now the only place one can find them is in books on the history of physics.

The true meaning of the law of electromagnetic induction was found by Maxwell. He also gave the law its simple and clear mathematical form based on the concept of a field that is today accepted by all.

Let us try to visualize the chain of Maxwell's reasoning and how he was able to perceive in electromagnetic induction a new fundamental property of an electromagnetic field.

Take an ordinary transformer. Connecting the primary winding to a house outlet, we immediately get current in the adjacent secondary circuit if it is closed. The electrons in the wire of the winding begin to move. But electrons have no "knowledge" of electromagnetic induction. So what forces bring the electrons into motion?

The magnetic field that links the coil cannot do it, because a magnetic field only acts on moving charges (that is how it differs from an electric field), and the conductor containing the electrons is stationary. * Then what is it that acts?

We know that in addition to the magnetic field there is an electric field that acts on the charges. And this latter field is capable of acting on stationary charges. That is its principal property. But the field we had in mind (an electrostatic field) is generated directly by electric charges, while induced current originates due to the action of an alternating magnetic field. Perhaps utterly new physical fields are involved, since the idea of close-range action is considered to hold.

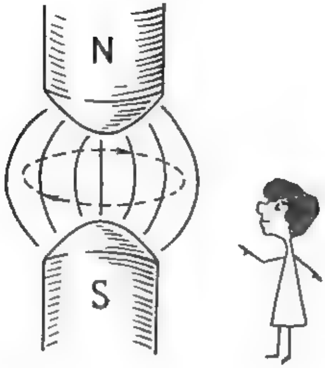
We shall not hurry with conclusions and seek salvation in conjuring up new fields, which was exactly what was done long ago when fresh forces were invented every time a difficulty was encountered. We cannot be sure, of course, that all the basic properties of magnetic and electric fields are already known. Constant fields are involved in the laws of Coulomb and Ampere which contain the main information about field properties. But suppose alternating fields begin to exhibit new properties? We shall hope that the idea of the unity of electric and magnetic fields, which has been so fruitful so far, will continue to serve us.

Then the only thing to presume is that the electrons are accelerated in the secondary winding *by the electric field and that this field is generated by the alternating magnetic field directly in empty space*. We thus have a new and fundamental property of magnetic fields: *a magnetic field varying in time generates about it an electric field.*

Now the phenomenon of electromagnetic induction appears in an entirely new light. The main thing is the process which occurs in empty space—the generation of an electric field by a magnetic field. Whether there is a conducting circuit (coil) or not is immaterial. A conductor

---

* Actually, electrons are in random motion in a stationary conductor, but the average velocity is zero since all the velocities cancel out. Hence the current induced by a magnetic field is zero.



with its supply of free electrons is simply an indicator (recorder) of the generated electric field, which sets the electrons in the conductor into motion and thus reveals itself.

The essence of electromagnetic induction does not at all lie in the appearance of an induced current, but in the generation of an electric field.

A rotational electric field

The electric field produced by a varying magnetic field is structurally quite different from an electrostatic field. It is not directly associated with electric charges and its lines of force cannot begin and end on them. In general, there is no beginning or end, they are closed lines like the lines of force of a magnetic field. It is the so-called rotational field.

Variations in the field of a strong electromagnet give rise to powerful eddies of an electric field which may be used to accelerate electrons to velocities close to that of light. This is the underlying principle of the betatron, an accelerator of electrons. In the betatron, an electric current is generated in a vacuum chamber without any metal conductors at all.

One may ask why this field is called electric? Particularly since it is of a different origin and has a different

configuration than a static electric field. The answer is simple: a rotational field acts on a charge in exactly the same way as an electrostatic field, and we believe this to be the main property of the field.

Another question that comes to mind is that since all the foregoing is actually only supposition and in no way self-evident, might it not easily be that matters stand quite differently? All the more so since we do not perceive an electric field directly and judge of its existence only from the forces operating on charged particles!

But this is essentially the old doubts about the existence of fields as such expressed by those who supported action at a distance. They are definitively dispelled by the existence of electromagnetic waves, in the very origination of which a fundamental role is played by the generation of an electric field via an alternating magnetic field.

Not all questions make sense

An alternating magnetic field gives rise to the eddies of an electric field. That may be. But doesn't this statement seem to be lacking in something? One obviously wishes to know the *mechanism* of the process. Is it not possible to explain *how* the connection between fields is actually realized in nature? This is where your curiosity cannot be satisfied, for there is no mechanism here at all. The law of electromagnetic induction is a fundamental law of nature (basic and primal that is). Its own action accounts for numberless phenomena, yet alone it remains unexplainable for the very simple reason that there are no laws more fundamental from which it could be deduced. At any rate, no such laws are presently known. Such is the fate of all fundamental laws: the law of gravitation, Coulomb's law, Ampere's law, etc.

We can, of course, ask all kinds of questions, but some of them have no meaning. For instance, one can and must investigate the causes of various phenomena, but there is no use in attempting to figure out why causality exists at all. Such is the nature of things and such is the world in which we live.

On symmetry

In electromagnetic induction Maxwell perceived the generation of an electric field by a magnetic field. The next and final step in the discovery of the basic properties of an electromagnetic field was made by Maxwell without any experimental evidence.

He was obviously guided by the same reasoning that prevents one from putting all the furniture on one side of the room leaving the other half empty. At the heart of matters lies the notion of symmetry, but not symmetry in a narrow geometrical sense, rather in a very broad sense.

The properties of symmetry are very fundamental to nature. That apparently is the reason why symmetry appears to us as a very necessary harmony of the world about us.

In electromagnetic phenomena it is not, of course, a matter of outward beauty and elegance in the things we observe with our sense organs. It is one of internal harmony that nature exhibits to the investigator of its most fundamental laws. Since man has an innate feeling for this harmony, he naturally seeks to find it even where the facts as yet do not fully display it.

A magnetic field gives rise to an electric field. Isn't there a reverse process in nature when an alternating electric field generates a magnetic field? This symmetry based reasoning formed the core of Maxwell's famous hypothesis about *displacement currents.*

Displacement currents

Maxwell made the assumption that such a reverse process actually occurs in nature. He gave the name displacement current to an alternating electric field in empty space or within a dielectric. It is called current because *this field generates a magnetic field in exactly the same way that an ordinary current does* (this is all the similarity there is between a displacement current and a conduction current). The modifier *displacement* indicates that it is no ordinary current, but a very specific type;

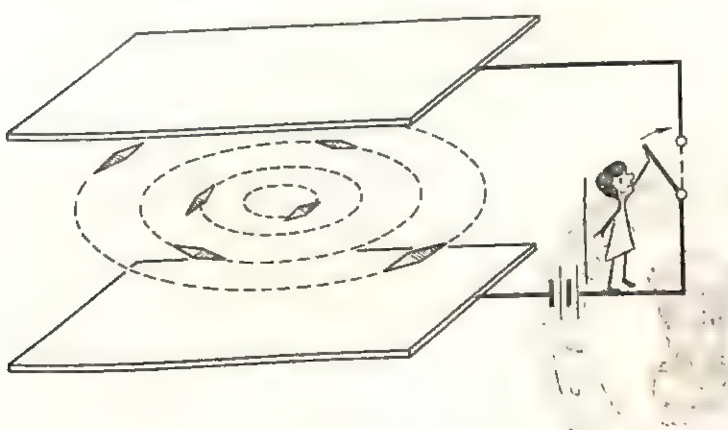
it likewise serves to remind us of that distant time when variations of an electric field in empty space were associated with the displacement of particles of the hypothetical ether.

For a long time Maxwell's assertion remained a hypothesis and nothing more. Yet there is every reason to call it a brilliant hypothesis, for its absolute truth was experimentally demonstrated.

Today it might appear rather commonplace. One might think that there is nothing unusual or striking in this conjecture. Perhaps any other scientist could have stated it? No, indeed. One should never forget that the very possibility of such a hypothesis arose only after electromagnetic induction was explained in terms of fields. And this was a time when most famous scientists were not taking the field concept seriously. There were several decades to go yet before any experimental proof of the existence of fields.

Maxwell not only advanced the hypothesis but straightway formulated a precise quantitative law that determines the magnitude of a magnetic field as a function of the rate of change of an electric field.

One is astonished by his consistent purposefulness, his confidence in his ideas when formulating the laws of the electromagnetic field. From the very start, when Maxwell undertook the study of electromagnetism following his successful work in the molecular-kinetic theory of matter, he made a point of reading only experimental papers on the subject, avoiding all theoretical discussions so as to have a free mind in the matter. This approach proved



remarkably fruitful and helped Maxwell to work out his own unencumbered point of view concerning electromagnetic processes.* Maxwell boldly placed at the heart of his quantitative theory an entity (the field) whose existence had not been demonstrated experimentally. Then, proceeding step by step, relying on experimentally established facts and regularities, he arrived at the final goal. The hypothesis concerning displacement currents was the final fundamental link. Here, Maxwell endowed his hypothetical entity with a fresh hypothetical property without direct experimental evidence (unlike preceding cases).

Operating in this manner, one might easily step out of science into science-fiction if the proper direction has not been taken. But this is not known at the beginning. It is precisely this proper choice of trend or direction in the construction of a theory that distinguishes a genius.

Thus we have yet another fundamental property of the electromagnetic field that cannot be broken down into more elemental components. In empty space an alternating electric field generates a magnetic field with closed lines of force (a rotational field). And in an increasing electric field the lines of force of the magnetic field form a right-hand screw with the field, unlike the left-hand screw for the field in electromagnetic induction. We shall see later on what a profound meaning lies behind this fact.

Proof of the reality of Maxwell's hypothesis lies in the existence of electromagnetic waves. The displacement current and electromagnetic induction fully determine the very possibility of their existence.

The electromagnetic field

After the discovery of the interrelationship between an electric field and a magnetic field, a very important fact came to light: these fields are not separate or in any way independent. *They are a manifestation of a single entity which may be termed an electromagnetic field.*

---

* We hesitate to recommend this method for general use because at that time an entirely new science was in the making—electrodynamics with its many specific peculiarities. And what is more, not everyone can be a Maxwell.

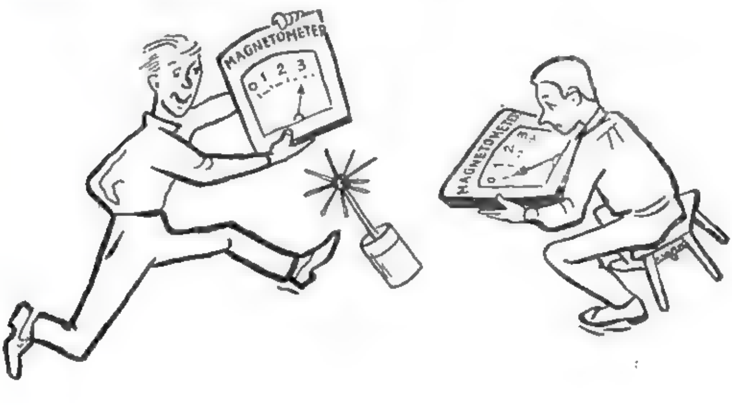
Let there be, in a certain area of space, an inhomogeneous electric field produced by some charge at rest relative to the earth. There is no magnetic field about the charge. But so it is only relative to the earth (or, in the parlance of physicists, in a frame of reference attached to the earth). Now for a moving observer an inhomogeneous field not varying with time will appear to be alternating. But an alternating electric field generates a magnetic field, and a moving observer will record a magnetic field in addition to the electric field.

In the same way, a magnet lying at rest on the earth creates a magnetic field, but an observer moving relative to the magnet will detect an electric field as well, in full accord with the phenomenon of electromagnetic induction.

Hence the statement that in a given point of space we have only an electric (or magnetic) field is meaningless. One must add: relative to a definite frame of reference. The absence of an electric field in a reference frame containing a magnet at rest does not at all mean that there is no electric field in general. This field may be detected with regard to any reference system in motion relative to the magnet.

Just like the colours of a landscape change when the latter is viewed through coloured glass, so the magnitude and configuration of fields vary as we pass from one frame of reference to another.

Just as blue objects become invisible when viewed through red glass, so a suitable choice of reference system



can in certain cases make a magnetic field unobservable.

There is however one very important circumstance. We can throw away our coloured glasses and say "these are the true colours of the landscape and this is what it is in reality". And one of the colour filters (the atmosphere) may be put in a privileged position. But this cannot be done with a frame of reference, for all of them have the same right to exist. For this reason, there is no configuration of fields of absolute significance, independent of any reference system.

8

Bacon on the laws of nature

The fundamental laws of nature, which include Maxwell's laws of electromagnetism, are remarkable in the following respect: "They can yield more than is contained in the material of which they are obtained." That is precisely why science is possible. For if every law, like sausage, contained only what it is stuffed with (to rephrase the words of Kozma Prutkov), there would be as many laws as there are phenomena in nature, and in place of modern science we would have a limitless aggregate of facts and information about observed natural processes, but we would never be able to predict anything.

This fact has to do with the very essence of science and so was grasped long before the laws of mechanics were stated. The words quoted at the beginning of this section belong to the English philosopher Bacon and were pronounced before Newton's great work *The Mathematical Principles of Natural Philosophy* was written.

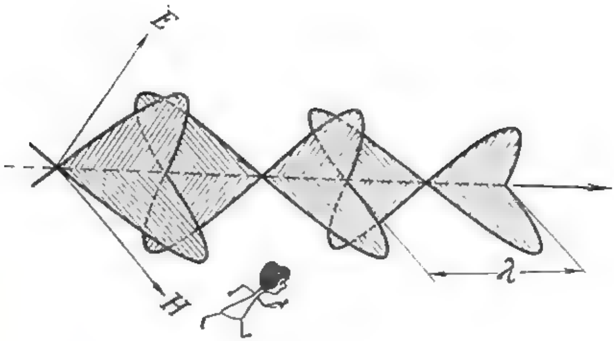
How electromagnetic interaction is transmitted

Among the numberless corollaries that follow from Maxwell's equations of an electromagnetic field, there is an extremely important result that would have been hard to guess beforehand. They contained, as Maxwell himself found, a finite velocity for the propagation of electromagnetic interactions.

On the action-at-a-distance conception, a Coulomb force acting on an electric charge will change instantly if a proximate charge is moved. Which is in keeping with action at a distance, since one charge "feels" the other directly through empty space.

Maxwell visualized this quite differently and in a much more complicated fashion. Any change of position of a charge alters the electric field in the vicinity. This changing electric field (displacement current) generates a varying magnetic field in nearby areas of space. But in turn a varying magnetic field will generate a varying electric field in accord with the field interpretation of events of electromagnetic induction, while an electric field in its turn will produce a magnetic field, and so forth. What is more, the originating eddies of a magnetic (or electric) field cancel out the field in areas where it already existed, but at the same time involve fresh areas of space. These events follow the rules for determining the direction of fields that we have already spoken about. If the fields were directed differently, this would upset the law of conservation of energy. A magnetic field created in space would build up in time all the while moving out in all directions.

Thus, the translation of a charge brings to life, as it were, the dormant potentialities of the electromagnetic field, and as a result, a splash of this field is transmitted involving ever greater areas of surrounding space and constantly rearranging the field that existed in its pathway prior to displacement of the charge. Finally, the splash reaches the second charge, which is what alters the force acting on it. But it will not occur at the very instant that the first charge changes position. The propagation of

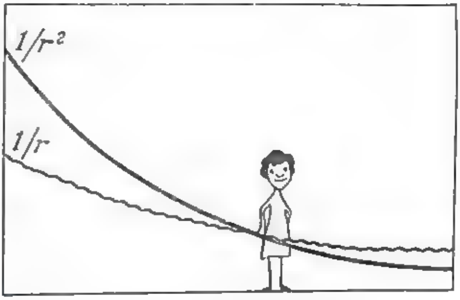


an electromagnetic disturbance, the mechanism of which Maxwell elucidated, takes place with a *finite*, though very great velocity.

How an electromagnetic wave originates

With pencil and paper and nothing more than the set of equations of an electromagnetic field before him, Maxwell was able to demonstrate, in pure mathematical form, that the rate of propagation of this process is equal to the velocity of light in vacuum, which is 300,000 kilometres a second. That is the new fundamental property of the field that finally makes it tangible reality. One could arrange an experiment to measure the time of propagation of a disturbance between two charges. Practically speaking, this would be very difficult to do because the velocity is so great. But that is not so essential. The significant thing was that for the first time it was possible to demonstrate experimentally the existence of a field. If the experiment is possible, a way will sooner or later be found to perform it. This is actually what happened when Hertz succeeded in obtaining electromagnetic waves.

Imagine that an electric charge was not simply moved from one point to another, but was brought into rapid oscillation along a certain straight line so that it moved like a weight suspended from a spring, but much faster. Then the electric field in the immediate vicinity of the



charge would begin to vary periodically. And the period would obviously be equal to the period of oscillation of the charge. The electric field would give rise to a periodically varying magnetic field, while the latter in turn would generate a varying electric field, which would now be some distance from the charge, and so on.

In the space around the charge, there would arise and move out to ever greater distances a system of periodically varying electric and magnetic fields, which would be propagated with the velocity of light. The accompanying illustration is probably quite familiar to many. This is what is called an electromagnetic wave moving out in all directions from an oscillating charge. At every point in space, electric and magnetic fields vary periodically in time, but since the farther away the point is from the charge, the longer it takes for the field oscillations to reach it, oscillations do not occur synchronously at different distances from the charge.

Maxwell was profoundly convinced of the reality of electromagnetic waves, but did not live to see them detected. He died rather young and 10 years before Hertz first demonstrated experimentally the existence of electromagnetic waves.

Interaction by means of electromagnetic waves

An entirely new type of interaction between electric charges is accomplished by electromagnetic waves. Waves

are emitted by oscillating electric charges, hence by charges whose velocities vary with time, or in other words which are undergoing *acceleration*. Acceleration is the chief prerequisite for the generation of electromagnetic waves. An electromagnetic field does not only radiate when a charge is in oscillation, but in the case of any sudden change in its velocity.

Thus, *the forces of interaction brought into play by an electromagnetic field depend not only on the distance between the particles and their velocities, but on their accelerations as well!* However, only the magnitude of the field depends on the acceleration. The force of the electric field of the electromagnetic wave acting on the charge depends, as before, only on the field intensity, but the force of the magnetic field depends also on the velocity of the charge.

The greater the frequency of oscillation of the charge, the higher the acceleration it will have and, accordingly, the more intensive the waves it emits. An increase of only twice in the frequency produces a 16-fold increase in the energy emitted. That is why radio-station antennas are oscillated with frequencies in hundreds of millions per second.

The most important fact of interaction by means of electromagnetic waves—and this is decisive—is the slow decrease in field intensity in the wave with distance from the source. We recall that electrostatic forces and the forces of current interaction are proportional to the inverse square of the distance and are considered long-range forces. Now in an electromagnetic wave, the decrease is inversely proportional to the *distance itself*! This is an extremely slow fall-off. All other forces diminish with distance much more rapidly. Also, as calculations show, the fields are found capable of moving away from the source to great distances due to successive excitation of each other. That is why the fields of even low-power radio stations may be detected at distances up to thousands of kilometres, while a static field would never be detected at such distances.

For this very same reason, we are able to see (light is also an electromagnetic wave) stellar clusters that are so far away that light takes hundreds of millions of years to reach us!

We must not forget about yet another aspect of the process of radiation. If a particle emits radiations, the outgoing electromagnetic waves carry off energy. The radiating particle loses energy and therefore must experience a certain retardation. It experiences something like a force of friction. But what is this force like? And how does it act?

We know that a charged particle is acted upon by a force from electric and magnetic fields. So far we have had in view only external fields, that is, those created by charged bodies around the particle. But there are still fields created by the particle itself. Do they exert any effect on the "source" that generates them? It is easy to see that so long as the "source" is at rest, there are no "self-acting" forces, otherwise we would have such an incredible thing as self-acceleration of a particle left to itself. The situation is also the same in the case of uniform and rectilinear motion of the source (which is obvious when we recall that rest is simply a particular instance of uniform rectilinear motion). In these most elementary cases, the "field tail scarf" tails along with the particle without pulling away or getting deformed in any way.

The picture changes, however, when we give the source a push, as it were. If the rate of propagation of electromagnetic signals were infinitely great, the field generated by a particle would follow just as close with every burst of speed of the particle and the force of self-action would remain the same—zero. But what we find is something different. The particle jumps out of its position of equilibrium in its own field; as a result, a force should appear tending to return it to the original position, a retarding force. The particle gets caught, as it were, in its own field, thus justifying the physicists' figure of speech: "radiation friction". We will not err by saying that the energy lost by a radiating particle is equal to the work done by the force of radiation friction—the force with which the field produced by the source acts on the source.

Now this self-action has yet another interesting feature. We have said that the force of self-action of a particle at rest (or in uniform rectilinear motion) is equal to zero. But it by no means follows from this that

the "energy of self-action" is also zero. The "field tail scarf" has energy and it has mass, and hence it makes its contribution to the inertia of the particle.

If an electron were suddenly to lose its charge (for some mysterious reason), its mass would instantly be reduced. By how much? We don't know yet. Which is not surprising, since this touches on aspects of particle and field interaction that require a profound knowledge of what is often called the structure of elementary particles. That is the science of tomorrow.

CHAPTER WITH NO NUMBER

In the overt world, it is you!
All secrets hidden away are you!
And no matter where I turn—
again it is you!

Book of Wisdom (Jami)

ELECTROMAGNETIC FORCES IN ACTION

1 *How Do Electromagnetic Forces Manifest Themselves?*

2 *Forces, the Structure of Matter, the Equations*

3 *Electromagnetic Forces in Electrically Neutral Bodies*

4 *Free Charges and Currents in Nature*

5 *Electromagnetic Waves in Nature*

6 *Why Electromagnetic Interactions Take Up Most of This Book*

7 *An Insertion with All the Rights of a Real Chapter*

1

If the reader has
any objections—

We realize that a chapter with no number is rather risky. But there doesn't seem to be any other way out, since we are not going to introduce any new forces and will continue to discuss electromagnetic interactions. But while in the preceding chapter we talked about fundamental problems dealing with the very nature of electromagnetic forces, here we shall attempt to show how a few, actually "basic" laws of electromagnetic fields enable us to grasp, at a single stroke, an enormous range of phenomena covering everything from the simplest everyday events to truly magnificent occurrences.

That is why there is no new number to this chapter: we continue the same story of electromagnetic forces but in a new way.

Electromagnetic forces and how they make themselves known

We might start out with the question: "Are there such things in nature as hidden manifestations of electromagnetic forces?" The answer is definitely this: most of the time we deal with implicit manifestations of these forces, though of course everyone of us has surely seen some real obvious displays.

Positive and negative charges or, to be more exact, positive and negative charged particles are, with rare exceptions, *bound* together forming neutral bodies. And the bond is usually deep inside matter, in atoms. Only here is the direct interaction of Coulomb forces decisive. But it is so deep and hidden that only complex physical apparatus is able to detect it. Most of the time we encounter electromagnetic interactions between neutral systems (atoms and molecules). This is interaction of *bound* charges, in which electromagnetic forces do not operate in the simple form of the laws of Coulomb and Ampere. So let us call this manifestation of electromagnetic forces implicit. Charged particles in the free state are much more rare than in the *bound** state. Cases in nature when charged bodies interact in accord with Coulomb's law or currents follow Ampere's law are comparatively rare. No wonder then that for centuries people lived with forces of an electromagnetic nature without suspecting them in the least. It never occurred to anyone that elasticity, friction and the like are simply different expressions for fundamentally the same forces.

When electromagnetic forces become short-range forces

The point is that electromagnetic forces between charges in neutral systems are *short-range*. They diminish with distance much faster than Coulomb or Newtonian forces. For this reason, these forces are apparent only at short distances, actually when the bodies are in contact. Here is disguised the fact that in reality the interaction between the bodies is always accomplished at a distance

* *Free* and *bound* charges are, incidentally, the actual terms.

and via the electromagnetic field and there is no direct contact at all.

Apparent interactions, though discovered long ago and observed only in special conditions, appeared as an oddity quite disconnected from everyday things. It was obvious that these forces acted through empty space without any contact.

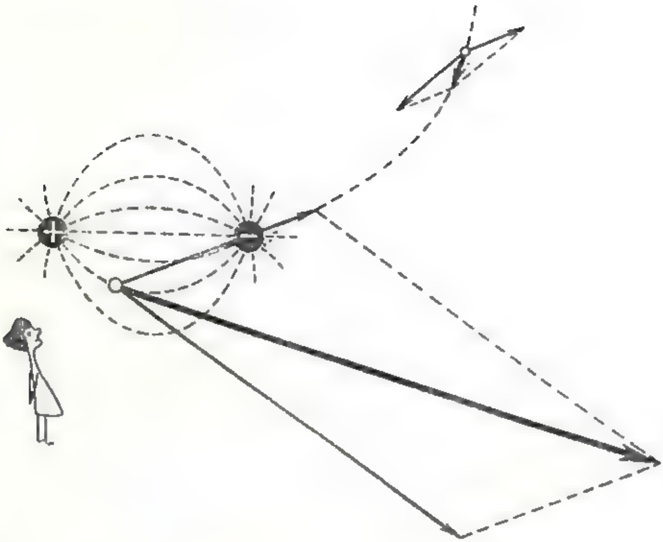
Oppositely charged particles form bound states by themselves and cease to exert any perceptible effect even on nearby bodies. Only the very closest neighbours are accorded any attention. But charges in such states are not able to move in an electromagnetic field independently of one another and cannot form an electric current of conduction.

This is the effect achieved by justice in bygone times when criminals were chained together in couples. They could neither run away nor do any harm to others about them. As Pushkin says in one of his poems,

> "They caught us, and smithies
> Forged the chain that bound us,—

In the atom, the chain that binds the electrons to the nucleus is in the form of an electric field.

Our comparison does not of course give any clue to why neutral bodies do not in the ordinary state exhibit elec-



tromagnetic interactions over any considerable distance. This is not a complicated matter in itself. Take the simplest possible neutral system—an electric dipole. Two identical charges of opposite sign a short distance apart. In a point at a distance from the dipole that is much greater than the length of the dipole, the electric fields of positive and negative charges are practically the same and in just about opposite directions. The total field (the sum of the two fields) is thus extremely small. In a dipole, the electric field decreases in proportion to the inverse cube of the distances, and even faster in more complex neutral systems.

Put differently, nearly the entire electric field is concentrated between the charges; the lines of force lie between the charges, pulling them together, as it were. We can say that there is no* electric field at large distances from the dipole. The whole field is then concentrated within the neutral body and on the very fringe of it.

---

* This holds for a static dipole whose charges are not displaced relative to one another. If the charges are in rapid oscillation, the dipole will radiate electromagnetic waves.

It is not always wise to begin with the simple

In our story about the action of electromagnetic forces in nature we shall nearly always be dealing with its hidden manifestations. Such cases as lightning or the discharge of torpedo fish or St. Elmo's light, and so forth are all interesting and exciting enough, but cannot compare in any way to the importance of such phenomena as elasticity, friction and the like.

We might begin with cases of *free* electric charges settling on bodies or moving between bodies. But free electricity arises out of bound electricity and does not exist in nature for a long time in that state (if we disregard states of matter at very high temperatures*). And so to reach lightning it is best to start from bound charges in neutral bodies. That is what we shall have to do, though hidden manifestations of electromagnetic forces are much more complex than apparent forces. Otherwise we would only learn how separate simple electromagnetic processes *develop*, and would be in the dark about why they *appear* and why they do not exist for infinitely long periods of time.

It might not have been necessary to go into so much detail if it weren't for one circumstance. In electrical engineering, which is where modern man gets most of his information about electromagnetic forces, we deal mainly with free charges, and those mostly in motion—electric current. Man has to this day been unable to harness electrostatic forces, though they are far more powerful than magnetic forces. Whence one often gets a distorted picture of the relative significance of the various forces in the world about us.

Nature is much more economical in the use of electromagnetic forces due to the fact that it preferres electric (Coulomb) forces as the more powerful and reduces the role of magnetic forces on earth to a bare minimum. Nature has turned out a more ingenious engineer than the human species. We shall attempt to show how nature

---

* At very high temperatures, a substance turns into plasma which is justly called the fourth state of matter, the first three being solid, liquid and gaseous.

does it. We shall try not to touch on the technical applications of the laws of electrodynamics because our problem is primarily a discussion of the *forces of nature.*

2

What one has to know to explain elasticity

Let us go back to our table at the beginning of chapter three. How is one to account for the fact that when the table bends, a force arises tending to return the table to its original state? We have already said that this force is of an electromagnetic nature. Now, since we know the principal laws of electromagnetism, we ought to be ready to explain the origin of this elastic force.

But just try. Nothing will come out of it. And not because we have no experience in the matter and we don't know where to begin. Knowledge of the nature of electromagnetic forces alone is not enough by far. To theoretize without facts is dangerous, for then a person begins to pressure the facts to fit the theory instead of using the facts to support it. That is how Sherlock Holmes expressed it. And it goes both for untangling a crime and investigating nature.

So what do we need to know in order to get a clear picture of the origin of elastic forces?

The laws of electromagnetic interactions enable us to state what forces arise between *charged* particles at *definite* distances apart, if they are moving at known *velocities*. To find the magnitude of these forces, one must therefore know (in addition to the fundamental laws of interaction) the *properties of the particles* that make up the substance, their mutual *arrangement*, and the *motions* they have. Without this information, we shall

not be able to explain the origin of elastic forces or of the forces of friction or, in fact, any other forces of an electromagnetic nature. Neither will we be able to understand why solids tend to retain their form and liquids their volume.

The three elephants that hold up physics

There is nothing at all surprising here. The elastic properties of, say, rubber only faintly resemble the properties of a wooden stick, though in both cases the elasticity is of an electromagnetic nature. This and similar facts are explainable only due to differences in the structure of matter. We mentioned structure when we discussed hidden and obvious manifestations of electromagnetic forces. But there will be more to say later on.

Suppose we know the structure of a substance. Is that sufficient to account for elastic and other forces of an electromagnetic origin, and does that permit us to understand why pieces of the substance are stable? In compression and tension, bodies change their dimensions and, hence, the distances between the charged particles that compose them. A change occurs in the state of the motions of the particles, their velocities, their translations. In order to find out how a particle will move under the action of some force—and in the theory of elasticity we must know this—we have to be acquainted with the laws of *motion*: how this force affects the motion. To account for the stability of pieces of a substance we must also know the equations of motion, because a substance is composed of moving and interacting particles, and it is precisely due to this motion that we have stability both of the atom itself and of macroscopic bodies, aggregates of enormous numbers of atoms.

We are familiar with the classical equations of motion, Newton's laws. These equations, together with the law of universal gravitation, enable us to explain the planetary motions of the solar system and, at the present time, to make extremely accurate calculations of the flight-paths of spaceships. A knowledge of gravitational forces alone is not at all sufficient.

To sum up, then, one has to know the *fundamental laws of interaction* and also the *structure of matter* and the *equations of motion*. Incidentally, all this is needed to account for any physical phenomenon. The structure of the substance, the forces involved, and the equations of motion are the three elephants on which physics rests.

The structure of matter

We already know what forces and equations of motions are. Ideas about the structure of a substance include, first of all, a knowledge of the properties of elementary particles. Information about the principal stable combinations that these particles form (atomic nuclei and atoms) may also be included in structure. A knowledge of structure likewise includes information about ordered formations of atoms—molecules and crystals. The latter is particularly important in accounting for the forces of elasticity.

Don't try to encompass the unencompassable

This is quite a pickle we've got ourselves into. We started out with forces, and now we find we have to discuss the structure of matter and the equations of motion. But that is the whole range of physics! There didn't seem to be any special difficulties when we dealt with gravitational forces. These are forces which become obvious only for large bodies, the internal structure of which has no effect at all on the magnitude of the force (only the masses are significant). The laws of motion are simple and obvious, they are Newton's equations.

It is quite a different matter when we come to electromagnetic forces inside neutral bodies. Here we require information about the properties of elementary particles, about the structure of atoms, molecules and crystals. And the most important thing is that the motions of atomic particles, the interaction between which ultimately accounts for the stability and elastic properties

of the substance, obey far more involved laws of motion than the classical laws. These are the laws of quantum mechanics, which need a whole book of explanation.

We have to confine ourselves to the *forces* of nature and so we will not go into any other fields of physics in detail—only so far as is really necessary. We shall give only a very rough and simplified description of the behaviour of particles, for a true understanding of which one would have to make a thorough study of quantum mechanics. Otherwise we will find ourselves in the position of a person who starts out with an incident in his life and goes on to give his full biography for fear he will be misunderstood.

The unity of nature

Incidentally, the elephants on which physics rests are not really independent beings. The peculiarities of one shape the characters of the others. It is only in the theory of elementary particles that we do not detect as yet any organic relationship between the properties of the particles themselves, the forces, and the equations of motion. It is not at all clear as yet why we find only a definite number of elementary particles in the world and why they have precisely the properties that have been experimentally discovered. In this sense, the problem of the structure of elementary particles is not solved, though there appears to be some hope of resolving it in the near future, at least in part. Relationships are emerging, and we are becoming confident that only the limitations of our scientific viewpoint create the illusion of three independent pillars of the theory. There is more reason to believe that the edifice of science should rest securely on one tortoise, called the *unified field theory*. What we term "independent elephants" are in reality probably segments of the tortoise's back.

The structure of atoms, molecules and macroscopic pieces of matter is fully defined by the known forces of interaction between the particles that make up these entities, and by the laws of their motion. We also have to know what particles make up our atoms and substance

Again, only experiment can give us this information. The remaining structure of atoms, molecules and so forth can in principle be obtained by means of pencil and paper. True, in many cases only in principle. The difficulties are so great, particularly when the system consists of a large number of particles, that the basic structural facts still have to be extracted experimentally.

As a rule, investigators penetrate the secrets of structure by means of direct experiments before they are able to study structure on the basis solely of the fundamental laws of interaction and the equations of motion.

The latter can frequently be stated when we pose the problem of explaining known facts about the structure of matter.* For instance, at the present time we know what the atomic nuclei of all elements are made up of, but we do not have a definitive theory of nuclear forces and so are unable to predict, theoretically and with full confidence, how stable a given combination of protons and neutrons will be.

The simplest way

Quite obviously, it is much simpler to explain by means of known forces and laws of motion the experimentally established facts of structure than to attempt, via these laws, to find out how a substance should be constructed. Just as it is easier to figure out the design of an automobile and understand how it works than to design and build a new one with the list of parts and engineering science to go by. No wonder there are patents to protect the rights of inventors.

We shall take the easier way and consider that the basic information about the structure of atoms, molecules and macroscopic bodies has been established experimentally, without saying how this was done. Our aim will be to relate how one can account for this structure on the basis of the action of electromagnetic forces. After that we can see what happens to a substance in the case of external effects. What forces are operative in it and why?

---

* The law of universal gravitation was discovered with the aid of the experimentally established laws of Kepler.

The atom

"Give me a set of electrons and Coulomb forces and I'll build an atom," said the atomic nucleus. If that's what he said, he was right, because Coulomb forces are the ones that hold electrons around a positively charged nucleus. If all the electrons are torn away from the nucleus, the electric field of the nucleus will immediately begin to capture any free electrons passing by in the vicinity and will go on capturing them until the number of electrons becomes equal to the nuclear charge. As soon as the electron-nuclear system becomes neutral, we have a finished atom.

Basic atomic structure is familiar to all. In the centre of the atom is the nucleus which contains almost the entire mass of the atom. Electrons form the outer structure of the atom, circling round the nucleus.

An atom is small and extremely empty inside if one disregards the electric field that permeates all of this space. It is much emptier than our solar system, which is hundreds of times the size of the sun and tens of thousands of times the size of the planets. If an atom grew suddenly to the size of the earth's orbit, the nucleus would be a thousand times smaller than the sun. Now if our sun were to diminish in size one thousand times, we would see only a bright point in the sky instead of the brilliant disk it is.

One often hears it said (and in earlier years that was the firm opinion of all) that electrons revolve about their nucleus in definite orbits like the planetary orbits of the solar system. Electrostatic forces are quite similar in character to the forces of gravitation, aren't they? The only difference is that the force of interaction of the "planets" of an atomic system (electrons) with one another

does not differ greatly from that with the nucleus, whereas in the solar system only the sun's attraction is great. Planetary interaction makes only a small contribution. The charge of the heaviest nucleus does not exceed the electron charge by more than a factor of 100, whereas the mass of the sun is a thousand times more than the combined planetary mass. And, of course, electrons repulse, while the planets attract.

But it is not this that makes the colossal difference between the structure of atoms and that of the solar system.

We are still in the dark about the origin of the solar system and there are probably reasons for the specific sizes of the planetary orbits. We can readily assume that they might have been quite different because man is constantly generating minute planets with all kinds of orbits, depending on the velocity of the rocket.

In the atom the situation is radically different. Its properties, and hence its structure, are utterly independent of its origin. All the atoms of a given chemical element are identical irrespective of whether they have existed since time immemorial or have come to life before our very eyes when a freshly produced nucleus captures some electrons. What is more, we cannot make an electron move in an atom any way we would like.

The whole point is that a nucleus builds its atom by means of its electric field and does not follow the rules of Newtonian mechanics in doing so, nor does it follow the rules of Maxwell's electrodynamics. It is impossible to build an atom on the basis of these laws.

The electrons in an atom cannot move in a straight line, naturally. They move with acceleration and consequently must radiate electromagnetic waves. Radiation is accompanied by loss of energy and therefore the electrons must inevitably fall onto the nucleus just like a satellite moving through the upper atmosphere loses energy due to air resistance and finally falls to earth. The difference is that a satellite can go on revolving for years, while an electron, if one follows classical theory, should come down in less than a millionth of a second. A momentary flash of light would indicate the demise of the atom. The electromagnetic field of classical physics should kill the atom, though it creates the atom too.

Something like Taras Bulba in Gogol's story: "I created you and I shall kill you."

Actually nothing of the kind happens. If the atom is not handled roughly, it will continue existing for any length of time. Nature did a wise thing in subjecting the motion of microparticles to quantum laws. Atoms, like the drivers of automobiles stay alive by following the rules—quantum rules for atoms, traffic rules for drivers. The only difference is that the quantum rules of intra-atomic traffic are laws of nature that not a single electron or anything else could infringe even if it wanted to. The gist of these rules consists in the fact that the energy of an electron in an atom can have only very specific, discrete values, it cannot vary continuously, and an electron cannot radiate continuously. There is always a minimal energy which an atom does not lose under any conditions as long as it succeeds in retaining its electron complement.

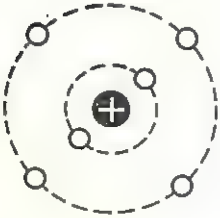
Radiation will be discussed later on. A very important thing to bear in mind all along is that the motion of an electron in an atom has hardly anything at all in common with the motion of the planets in their orbits. If it were possible to photograph a hydrogen atom in its lowest energy state (the most elementary system), we would see a cloud with maximum density a certain distance from the nucleus. This distance could be taken for a rough measure of the radius of the orbit. This snapshot of the atom would be quite different from a picture of the solar system and would be more reminiscent of a smeared out spot that one gets when taking a long-exposure photograph of a moth circling at random round a lamp.*

There is one thing we must remember about the structure of complex atoms, and that is that the electrons are arranged in layers or shells. Every shell has a very definite number of vacancies. The innermost shell can accommodate only two electrons, the next one, eight, and so forth. The farther away from the nucleus a shell is, the more electrons it can accommodate, but the number is always limited. This limitation has nothing to do with electric forces but is completely determined

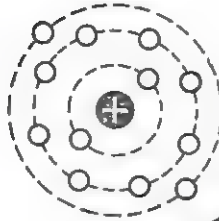
---

* The similarity is actually only superficial because electron motion differs fundamentally from that of any macroscopic body.

Carbon



Silicon



by rigorous quantum-mechanical rules. It is called Pauli's principle, which states that no two electrons can occupy the same state at one time. There must be some difference, says nature.

Increasing the number of electrons in an atom and the formation of new shells filled with electrons does not expand the atom. An increase in the positive charge of the atomic nucleus causes a compression of the inner shells. In this way, the dimensions of all atoms as defined by the radii of their outermost shells come out about the same, while the inner electrons become packed in closer to the nucleus as the charge builds up.

These regularities in atomic structure are most obvious when atoms make encounters. Then they impinge with their outer shells, nd what occurs deep inside the atoms is not very important. The significant thing is how many electrons there are on the periphery of the atom. Actually, their number completely determines the "intentions" upon collision: to join up or each go its own way. We can safely say that everything depends on the way they are "dressed", though the dress is determined by the core of the atom, the nucleus.

The number of outer electrons varies periodically as the charge on the nucleus increases. After completion of the first shell, a fresh shell is started. That is the key to the physical meaning of Mendeleyev's Periodic Table

of the elements. The point is that the chemical properties of an atom are determined by the number of the outer (loose-bound) electrons.

It will readily be seen that the fewer electrons there are in the outer layer, the more weakly they are bound to the nucleus. Roughly speaking, the inner electrons together with the nucleus may be regarded as a positive ion. If the outer shell has only one electron (typical of metals like lithium, sodium, etc.), it is attracted by an ion charge of one unit (in the atomic system of units).

When there are two electrons in the outer shell (beryllium, calcium, and so on), each of them is attracted to the centre with a force twice as great, because the electric charge of the remainder of the atom is equal to two, and so forth. As the number of outer electrons increases, so does the charge of the positive ion; the force of attraction of the electrons increases, the "orbital radius" diminishes, and the binding strength rises. The bond is the strongest when the outer shell is completely filled. This occurs in the inert gases helium, neon, argon and the others. Helium has two electrons in the outer shell, and the others have eight.

The atomic nucleus

In the atom, electric forces play the dominant role. Inside the nucleus, they are important but not dominant. The positively charged protons of the nucleus are packed in very tight and for this reason cannot help interacting. Their repulsive force is tremendous and the nucleus would not stay whole for an instant and would fly off in all directions with velocities close to that of light if it weren't for the far more powerful nuclear forces.

Powerful Coulomb forces of repulsion make the nucleus like a compressed spring ready to straighten out. In the atoms of the heavier elements there are so many protons (uranium has 92) that the nuclei become unstable. The nuclear forces of attraction, which in the light nuclei easily overcome electric repulsion, can hardly cope with the situation in uranium. Just the slightest push (an impinging neutron, for instance) is needed to make the

nucleus fall into two pieces that fly off at high speed in opposite directions under the repulsive forces. It is precisely these *electric forces* that produce the energy in an atomic reactor or in an atomic bomb explosion. The so-called nuclear energy that is released here is actually electromagnetic energy.

Two types of forces between atoms

It is not hard at all to prove the existence of considerable forces between neutral atoms (or molecules). Just try to break a big stick! It consists of atoms, as we know. There are two different types of electric force that operate between atoms. One force has a simple analogy in the interaction of gross bodies and is quite "comme il faut", for it is fundamentally classical.

The other type is a quantum-mechanical force, often called an exchange force. It is calculated by means of quantum mechanics and cannot be fully visualized because intra-atomic affairs are not amenable to descriptions in terms of classical physics, which is the only pictorial language for people whose everyday life consists in contemplating phenomena that obey classical physics.

At large distances, only classical forces are operative between atoms. The interaction then very largely disregards details of structure of the atoms themselves. Both the interaction of individual atoms and the interaction of groups of atoms combined into molecules obey the same law. This type of force is therefore called *molecular*. This is because an atom may be regarded as a particular case of a molecule, its most elementary form. Occasionally, these forces are termed Van der Waals forces after the Dutch scientist who first applied them in the theory of gases and used them to account for the transition of gases into the liquid state.

At appreciable distances, neither atoms nor molecules repulse. Distant neighbours always tend to come together. Molecular forces are attractive forces.

"Exchange" forces come into play when atoms come so close that their outer shells are in contact. This is where

the individuality of each pair becomes apparent. The atoms either form a stable system, a molecule, or are energetically repulsive.

The joining of atoms into molecules is already the realm of chemistry. For this reason, the quantum-mechanical forces of adhesion are frequently called chemical.

If we bring the atoms to distances less than the sum of their radii, repulsive forces begin to operate, and it is impossible to jam one atom into another.

Note that *electrostatic* interaction underlies both molecular and chemical forces. Magnetic forces do not play any appreciable role.

Let us examine these two types of force in a little more detail.

Molecular forces

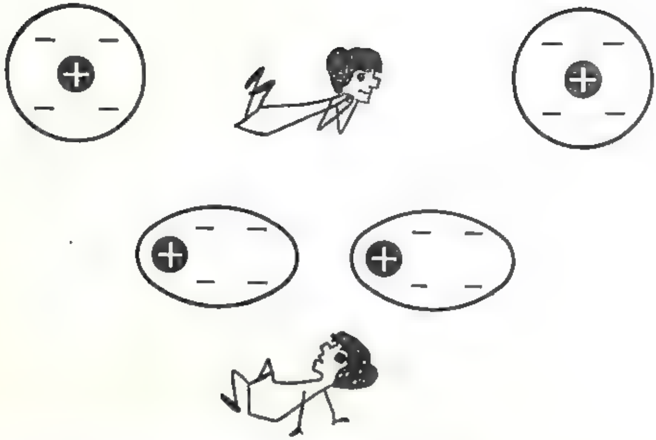
How does molecular attraction arise between electrically neutral systems? Let us first try to find out why pieces of paper and other light objects are attracted to an electrified body.

A positively charged rod is brought up to a strip of paper. The charged particles of the atoms of paper cannot remain indifferent. The electrons shift towards the positive charge, while the nuclei move back a bit. This is what physicists call polarization. A negative charge finds itself closer to the electrified body than a positive charge, and the force of attraction predominates over the force of repulsion.

If there were only a single molecule in place of the strip of paper, the very same thing would take place. The electric field "blows" the light electrons away from the nuclei, and the molecule turns into an electric dipole in which opposite charges are spatially separated.

In many substances, including water, the molecules are something like electric dipoles from their very birth. These molecules (via their electric fields) give rise to polarization of nearby molecules and to forces of attraction.

Attractive forces are absent only when the electron cloud of each atom has complete spherical symmetry.



Actually, however, only on the average for a large period of time can we say that the "centre of gravity" of the negative charge lies in the nucleus of an isolated atom. At a given instant, an electron (for the sake of simplicity we have in mind the hydrogen atom) can be detected anywhere within about 10^{-8} cm from the nucleus. When one atom approaches another, the electric field of the electron-nuclear system disturbs the electron motion of the adjacent atom so that the centre of gravity of the negative charge of the atom is displaced relative to the nucleus. Every atom (or molecule) polarizes its neighbour and they begin to attract one another.

This interaction is fundamentally Coulomb interaction. But since attraction between neutral systems is a consequence of a certain predominance over the existing repulsion, and since the polarization of the systems markedly loses strength with increasing distance, these forces are considerably weaker than pure Coulomb forces and fall off much faster with distance. If in an isolated molecule the positive and negative charges are not spatially separated, the attractive forces are inversely proportional not to the square but to the *seventh power of the distance*. When the distance increases twice, the force diminishes not four times but 128 times! That is why these forces are negligible beyond distances that exceed ten-fold the dimensions of the molecules. Van der Waals forces are short-range forces.

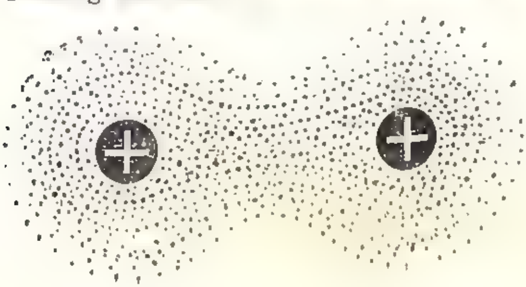
Chemical forces

Even Faraday was aware that chemical forces might be of an electromagnetic nature. He said that atoms of matter are somehow endowed with electric forces or are associated with them in some way and that this accounts for some of their most remarkable properties including their chemical affinity. Today the electric nature of chemical forces has been demonstrated rigorously.

Van der Waals forces are not able to account for the formation of molecules. First of all, they are too weak for this purpose. But that is not the most important thing. Chemical bonds, like friendship between humans, have a peculiar property of *saturation*. An atom of hydrogen can pick up one other such atom, but never two or three. An atom of carbon can link up four atoms of hydrogen, but no more. This property is mysterious from the very outset. So far not a single force has had this property of saturation. Like an orator with the spark of magnetism who draws crowds, a star can attract any number of planets. The force acting on one of them has nothing to do with the presence or absence of the others. Neither is there any saturation in electromagnetic forces between charged particles, nor in the forces of molecular attraction.

In chemistry, the property of saturation is expressed by the concept of valency, which was introduced long before scientists could undertake a study of the nature of chemical forces.

In very broad terms, chemical binding may be considered due to *collectivization* of the outer (valence) electrons of two combining atoms. At definite distances between the nuclei, the collectivized (common) electrons passing between the nuclei balance the repulsion of the



latter. At large distances there is no collectivization, and only Van der Waals forces are operative. Saturation is due to the limited number of collectivized electrons.

In the most elementary hydrogen molecule, both electrons behave as if each electron spent some of the time near one nucleus and the rest of the time near the other. It is precisely for this reason that the forces generated by electron collectivization are called *exchange forces*. However, exchange should not be taken too literally as the bandying about of electrons from proton to proton. There is no such pictorialness so typical of classical mechanics. The true meaning of the *exchange effect* consists in the simultaneous collectivization of two electrons by two identical nuclei.

The shape of the electron cloud of a hydrogen molecule (H_2) differs markedly from the spherically symmetrical cloud of isolated atoms. What we actually get is something rather reminiscent of a biological cell prior to complete division.

The atomic nuclei correspond to the nuclei of the daughter cells, the electron charge to the protoplasm.* In one case, the cells are held together by a strand of protoplasm until complete fission has occurred. In the molecule, this is done by an "electron protoplasm" which engenders a mutual attraction of nuclei by Coulomb forces, just as if a certain portion of negative electric charge were concentrated between them. If the distances between nuclei are not too small, the forces caused by electron collectivization fully balance the repulsion of the nuclei. For very small distances, the portion of charge concentrated between nuclei is not sufficient, and the electrons are as it were, squeezed out from between the nuclei into the peripheral region. Here the nuclear repulsive force is not balanced.

In this way, both attraction and repulsion are fully accounted for.

When unlike atoms join, the common electrons are in nonsymmetric motion relative to the two nuclei. This is particularly evident in the so-called heteropolar**

---

* From here on we have borrowed the vivid picture of the nature of chemical forces given by the Soviet physicist Ya. Frenkel.

** Unlike homopolar (uniform) molecules, such as the hydrogen molecule.

(nonuniform) molecules such as table salt (sodium chloride, NaCl), hydrochloric acid (HCl), and others. In sodium chloride, for instance, the binding is accomplished by the collectivization of eight valence electrons: one from sodium and seven from chlorine. Since the remainder charge of chlorine is greater, all common electrons are shifted drastically towards the chlorine ion, and the association is more in the nature of the "stronger" atom expropriating an electron from the "weaker" atom, which becomes a positive ion (the former becoming a negative ion), and the chemical binding reduces to the attraction of unlike charges.

The distribution of electrons loses its polar nature as the differences in the nuclei disappear, and becomes completely symmetrical for identical atoms.

A molecule should not be visualized as a sum of immutable atoms held in equilibrium by forces of attraction and repulsion. That was the crude and oversimplified way Berzelius pictured matters a hundred and fifty years ago.

There is no interatomic relationship in a molecule for the simple reason that there are no atoms that remain unchanged when they combine to form a molecule. For instance, there are no hydrogen atoms, strictly speaking, in a hydrogen molecule, because their individuality dissolves when they merge into a new system. The only thing that remains is the raw material out of which they are constructed: two protons and two electrons. Therein lies the fundamental difference of chemical forces from all other forces that have been discussed. A molecule may be regarded as the sum of nuclei (shielded by inner electrons) and common outer electrons, whose movements depend on the distance between the nuclei.

Spin of elementary particles

The last important question is to find out what determines the valency of an atom. This involves learning about one more entirely new and fundamental property of elementary particles, viz. *spin*. It was spin in conjunc-

tion with the Pauli principle that permitted Heitler and London (of England) to construct a quantum theory of chemical binding and explain valency.

Spin is essentially particle rotation, though, of course, it would be wrong to picture something like little tops spinning on their axes. Don't forget that particles are not tiny balls and in general are something beyond the powers of representation of the most imaginative painter. Our pictures are good for the gross world of big things (the macroworld), but are hopelessly inadequate for studying the microworld.

So spin is a spinning and yet the reader is warned not to picture any kind of mechanical rotation! That, you will say, is just a little bit too much. After all, then, what *is* spin?

Let us picture a bullet emerging from the barrel of a gun and revolving about its longitudinal axis (in the direction of flight). It strikes a target and imparts the rotational motion to the target. Now they both are turning on their axes in the same direction. Physicists say that the turning moment (torque) which the bullet originally had became distributed between the bullet and the target in which it lodged. The turning moment of a system of bodies cannot of itself, without external influence, either increase or decrease. This is due to the law of conservation of turning moment or angular momentum. It is not a matter of terms however. They are not so important for us. Neither will we need a precise statement of this most important conservation law, which justly occupies a place alongside the law of conservation of energy and momentum. The important thing is that one can judge (precisely, quantitatively)

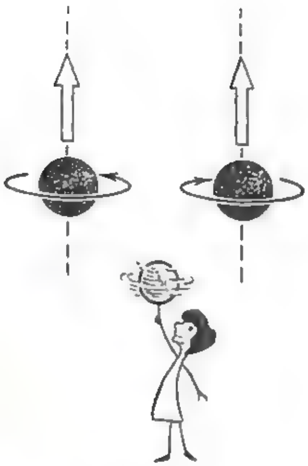
of the earlier rotation of the bullet from the rotation of the target.

Now imagine that a target is under bombardment by electrons or other elementary particles and absorbs them. If all particles are turning in one direction, they should impart to the target the same rotation upon being absorbed. Thus, the greater the spin, the stronger the *rotation of the target.*

There is no need for any fruitless search for an explanation of spin in terms of mechanical models. In target experiments we have a fundamental scheme of how this spin can be *measured.* That in itself is an achievement. So far we have of course only skimmed over the surface of extremely involved laws of motion and interaction of elementary particles... However, even this superficial glance enables us to compare the spins of various particles thus yielding some sort of picture of this new property of microworld entities.

Our target experiment was of course rather crude, but that fundamentally is the way things are done.

Getting back to our target, which is, say, a small coin, let us fire into it elementary particles all twisting in the same direction. A remarkable thing comes to light. For an equal number of hits, certain particles (electrons, protons, neutrons, and a number of others) transfer to the target the same angular momentum. Which means that their spin is the same. Photons (particles of light) transfer to the target twice that amount of angular momentum, while certain other particles such as pi-mesons,



for instance, do not cause any rotation at all. They have zero spin.

Quantitatively, the value of spin is known with great exactitude, for it is either 0, $\hbar/2$, or $\hbar$, where $\hbar$ is the famous Planck constant, or quantum of action, which we shall be coming across more frequently now. The electron has spin equal to $\hbar/2$. Planck's constant is so small (it has 27 zeros in the denominator) that our coin target will execute one rotation per second if we continue the bombardment for 10,000,000,000,000,000 years at the rate of 1,000 shots a second. Hardly worth talking about, one might say, but don't be in too much of a hurry. Our situation is something like an attempt to make the moon turn on its axis by firing pistol shots at it. Big or little spins have no meaning because the scales of the microworld are so different from our gross world. The significant thing is that spin is essential in many cases, one of which is when atoms combine to form molecules.

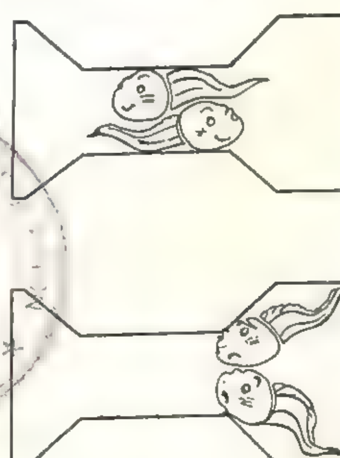
Experiments show that an electron can be twisted in only one of two ways: either rightwards or leftwards in the direction of motion. And the target is twisted accordingly. In other words, there are only two orientations of spin with respect to any direction. Thus, *if the spin of an electron is fixed, the spin of another one is either parallel or antiparallel to it.*

What determines the valency of atoms?

The mutual orientation of spins is a decisive factor in the formation of a molecule of hydrogen. Chemical binding appears only when the common electrons have opposite spins. When hydrogen atoms with parallel spins collide, the electrons also become common for a period of time, but no stable states arise. Collectivization leads to repulsive forces, irrespective of the distances between nuclei.

Quantum mechanics states that the mutual orientation of spins determines the type of motion of the electrons. In the case of antiparallel spins, the electrons spend a r latively long time between nuclei, so that the mean density of negative charge is sufficient to balance the repulsion of the nuclei. For parallel spins, this density is low and repulsion predominates.

Electrons with identically oriented spins cannot squeeze into the space between nuclei, like two tadpoles head to head trying to enter a tiny opening (see accompanying drawing). The internuclear space is big enough for electrons with opposite spins (antiparallel tadpoles fit it too). And that ends the analogy—electrons are really not tadpoles. Classical physics does not help us here in the least; it cannot explain why the orientation of spins is so



essential in affecting electron motion. The whole thing is a quantum effect.

And so a pair of electrons with antiparallel spins form the binding states.

Now we have all we need to account for saturation and valency. Let us take the most elementary case. Why can't a hydrogen molecule H_2 accommodate yet another atom? Why not have three electrons binding three nuclei?

That is forbidden by the Pauli principle. The common electrons are in the same quantum state and therefore have to differ as to spin orientation. But there are only two possible orientations! And so two electrons with antiparallel spins can accomplish a bond, while a third electron is one too many.

Once the hydrogen molecule has been formed, it will always repulse hydrogen atoms, thus accounting for saturation.

Now there is one more thing. Each of the hydrogen atoms merging to form a molecule has *one electron with an arbitrarily oriented spin*. The hydrogen molecule has *a pair of electrons with antiparallel spins* and does not attach fresh atoms.

This fact is of a very general significance. In each of the atoms, the electrons that form pairs with antiparallel spins do not participate in chemical binding. Such bonds are accomplished only by *electrons with "free" spins.*

In the inner completely filled shells, electrons always form pairs and never take part in chemical binding. The situation is exactly the same in the case of the noble-gas atoms, which for this reason are totally inactive chemically. Only when the outer shell of the atom is not completely filled can its electrons establish a chemical bond.

But not all electrons! *The number of electrons with "free" spins (and hence the valency of the atom) is equal either to the number of external electrons outside of filled shells or to the number of electrons that would fill up a shell completely depending on which of these two numbers is smaller.*

Thus, despite its exceedingly minute value, electron spin determines the whole chemistry of atoms. Chemical

reactions are involved in a tremendous range of events from simple burning to the most intricate transformations within a living organism, bringing about fundamental changes in the world about us.

Gases, liquids and solids

Let us try to delineate very roughly the structure of gases, liquids and solids. The molecules (or atoms) of a gas are in swift random motion throughout the volume occupied by the gas. The distances between particles greatly exceed their dimensions. Constant collisions at high speed send the particles off in all directions building up a pattern of quite random movements.

The molecule of a liquid behaves differently. Held tight between the other molecules, it swings about a position of equilibrium, unable to do more than "run on the spot". Only from time to time does it jump out of its cage into another cage formed by other molecules. It settles down for only about a ten-millionth of a second.

The atoms of solids are not able to tear away from their immediate neighbours and have to mark time where they are. True, very occasionally they, too, break out of their equilibrium position.

Let us not forget about yet another important difference between liquids and solids. To put it crudely, a liquid is a tightly squeezed pushing and pulling crowd of individuals without any order, whereas a solid body is, as a rule, a very orderly structure of individuals which, though not standing at attention due to their thermal motion, keep within a definite range of intervals. If we connected the centres of the equilibrium positions of the atoms or molecules, we would obtain a regular lattice (a space lattice of course, not a plane lattice), called a crystal lattice. Most solids are of crystalline structure.

It is only in amorphous bodies like glass that we fail to find any order in the arrangement of the molecules. For that reason, they are frequently not regarded as solids but rather as very viscous liquids almost completely devoid of fluidity.

The elastic properties of liquids and gases

Now we have enough information to grasp the origin of elastic forces in liquids and gases. The reader will probably be able to develop this picture himself. Try it, and then read about it here a little later. Since this won't be exciting, let us continue with the story about surface tension, which is a bit more complicated.

Of the attractive forces operative in gases and liquids (with the exception of liquid metals) we have only Van der Waals forces, while in solids there are exchange forces as well.

Van der Waals forces hold the molecules of a liquid close to one another at distances of the order of the dimensions of the molecules themselves. If a liquid is compressed, the molecules will come closer together and repulsive forces will build up very rapidly. As it is, the molecules are so close that just a slight decrease in distance will result in the development of tremendous repulsive forces. Something like getting into an overloaded bus.

It is not much harder to account for the fact that a liquid flows and is not able to retain its shape. Molecular jumps (which we have already mentioned) occur in the direction of an external force, say that of gravity, and the liquid begins to flow. However, the time during which the force acts must exceed the settled-life time of the molecules, otherwise the force will only produce an elastic shearing deformation, and ordinary water will be as hard as steel.

Upon heating, the energy of thermal motion of the molecules increases and molecular jumps become more frequent. Finally, the Van der Waals forces are no longer able to hold the raging molecules in place, and the liquid ceases to exist. Gas is formed.

The molecules of a gas fly about in all directions, for molecular attraction is no longer strong enough to handle them. The substance loses both its shape and its volume. No matter how big we make the container, the gas will fill it up completely with the greatest of ease.

The incessant beating of numberless gas molecules against the walls of the cotaining vessel produces pressure.

Surface tension

Such forces as gravity, elasticity and friction make themselves felt all the time. Now there is another common force that we hardly pay any attention to. It is not very strong and no great effects are ever produced. It isn't even required any more at the entrance examinations to institutes and universities. Yet we cannot pour water into a glass or do anything with a liquid, for that matter, without bringing this force into action. It goes by the name of surface tension.

We are so accustomed to the effects produced by surface tension that we do not pay any attention to them, except perhaps when making soap bubbles. Yet they play a rather important role in nature and in our everyday life. Surface tension permits us to write with ink. The old-fashioned pen would never pick up any ink, while the fountain-pen would unload its entire supply at one time. One wouldn't be able to soap one's hands, and no suds would form. A little rain would get one soaking wet, and rainbows would never shine. The water cycle of the soil would be disrupted and plants would wither away. Important functions of our own body would suffer as well.

The simplest way to observe the nature of surface tension forces is in the formation of a droplet of water from a loose faucet. Watch a drop grow, form a waist, and then drop off. Just a little imagination is needed to see that the droplet is contained in a tiny elastic bag, as it were, which breaks open when the drop gets too heavy. Actually, of course, there is nothing but water in the drop. It is the surface layer of water that behaves like a taut elastic film.

We get the same impression from the film of a soap bubble. It is so much like the taut rubber of a balloon. If you take the straw out of your mouth, the bubble will release the air inside and collapse.

Put a needle carefully on the surface of water in a dish. The surface film will sag but it will hold up the needle. That is what permits tiny water measurers to glide over the surface like skaters over ice.

Water poured carefully onto a fine-mesh sieve will not flow through due to the sagging of the surface film of water. So you can carry water in a sieve. Which again

shows how hard it is at times to say a real piece of nonsense. Fabrics are actually sieves formed by interlaced threads. Surface tension prevents the water from seeping through. That is why it takes some time to become wet "through and through".

If it weren't for gravity, surface tension would force a liquid to take on a spherical shape. The smaller the droplet, the greater the role of surface forces as against volume forces (gravitation). That is why tiny droplets of dew are almost spherical in shape. In free fall we have a state of weightlessness. For this reason, raindrops are nearly perfect spheres*. The sun's rays are refracted in these drops of rain and we get a rainbow. If the raindrops were not spherical we would not have any rainbows, says theory.

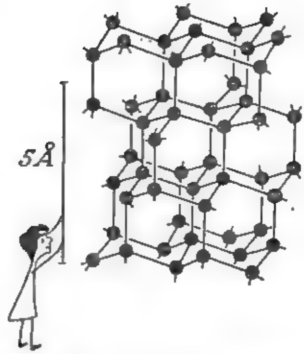
The forces of surface tension manifest themselves in such a multiplicity of ways that we cannot even begin to enumerate them. Our job will be to explain why they occur.

If a large group of entities have the property of attracting each other, they will collect like a swarm of bees, each one trying to get inside, thus making it into a sphere. That is a model of how surface tension originates.

Molecules of water (or any other liquid) attract with Van der Waals forces and form a collection of entities that tend to come together. Each molecule on the surface is attracted by all others and therefore tends to move into the interior both in liquid and solid bodies. But unlike solids, liquids possess the property of fluidity due to molecular jumps from one stable site to another. This enables a liquid to take up a shape in which the number of molecules on the surface is a minimum, which is the surface of a sphere (least surface area for a given volume). The surface of the liquid contracts and we recognize this as surface tension.

The origin of surface forces is quite different from the elastic forces of a stretched rubber film, it appears. When rubber contracts, the elastic force weakens, but the forces of surface tension do not change in the least

* Any deviation from sphericity is due to the resistance of the air. Inside a spaceship spilled water would stay in a perfect sphere.



as the surface film is reduced because the mean distance between molecules remains the same.

That is why the origin of surface forces cannot be explained in such a simple pictorial fashion as elastic forces, where a change in intermolecular distances is involved. Here the situation is more complicated because surface-tension forces appear in the course of a complex reorganization in the shape of the liquid without any change in its volume.

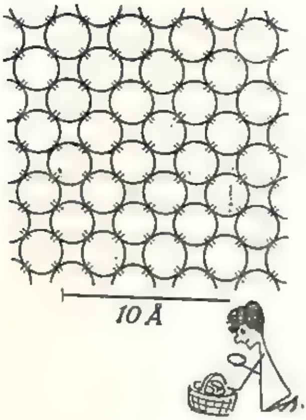
Four types of crystals

Diamonds and a paraffin candle. The former is a symbol of hardness, the latter (like wax) is soft and pliable. One might think that the opposite sets of properties are in accord with totally different combinations of the units that make up these substances.

That is quite right. Paraffin consists of separate molecules held together by Van der Waals forces. A crystal of diamond may be regarded as a single giant molecule. The forces of molecular attraction are very much weaker than the chemical forces and that is why paraffin is way behind diamond in hardness.

Crystals which consist of isolated molecules are called *molecular**. Diamond is a *valent crystal.*

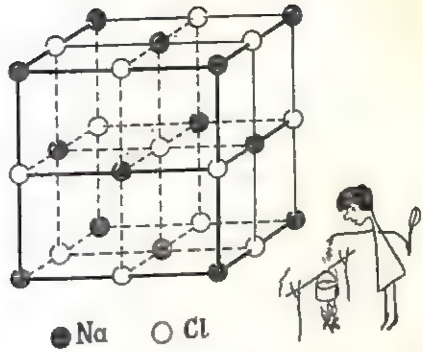
* Molecular crystals include those consisting of homopolar molecules: hydrogen, nitrogen and so forth. Dry ice (solid carbon dioxide) and many organic substances are also molecular crystals.



This name comes from the fact that each carbon atom in diamond has four neighbours, which is exactly equal to its valency. Any two neighbours establish between themselves an electron-pair linkage via a single electron. But do not think that the collectivized pair belongs only to two atoms. From each atom there are four pathways (bonds) leading to neighbouring atoms, and a given valence electron can move along any one of them. It can move from one atom to the next and go on wandering along the bond pathways throughout the whole crystal. We can give a flat picture of the crystal lattice of diamond in the form of rings packed tight against one another, while the electron-pair valence bonds are depicted as double dashes.

The only rule that must be scrupulously observed by the electrons is Pauli's principle: no more than two electrons can move along any pathway at one time. The common valence electrons belong to the entire crystal as a whole, which makes the crystal one stupendous molecule.

Electron-pair bonds of diamond are very strong and are not ruptured with increasing intensity of thermal oscillations of the atoms (that is to say, as the temperature rises). That is why diamond does not conduct electricity. The valence electrons which bind the atoms together are held firmly in their crystal lattice, and an external electric field does not exert any noticeable effect on their movements.



Crystals of silicon and germanium are much like diamond crystals, but their electron-pair bonds are not so strong. A slight rise in temperature ruptures some of the bonds. The electrons leave their trodden paths for freedom. If there is an external electric field, they move along the nodes of the lattice producing an electric current. These substances are called semiconductors.

Collectivization of valence electrons likewise binds the atoms of what are called *ionic crystals*: sodium chloride (NaCl), silver bromide (AgBr) and others. In the case of NaCl, as you remember, collectivization reduces to chlorine taking one electron of sodium. The same occurs in a crystal of table salt. All the valence electrons are actually in movement along the chlorine nodes of the lattice and thus, roughly speaking, the crystal consists of ions of opposite sign. Here the bond is established by electrostatic forces of attraction.

There is yet a fourth type of crystal—*metals* and *alloys*. When a piece of metal is formed out of separate atoms, the valence electrons lose all bonds with the atoms and become the property of the whole chunk of metal. Positive ions float in a negative liquid, as it were, formed by all the common electrons. This "liquid" fills all the interstices between the ions and pulls them together via Coulomb forces. The binding here is therefore of a chemical nature, as in the case of the valence crystals*.

But there is a big difference. In valence crystals the common electrons circulate along strictly indicated pathways

* The nature of bonds in liquid metals is the same as in solids.

between neighbouring atoms. Now in a metal the electrons are quite free and can move in any direction about the chunk. This is clearly evident from the fact that metals and alloys are good conductors of electricity, whereas valence crystals are for the most part insulators.

The relative freedom of valence electrons inside metals is due to the fact that the bond is very weak between electrons and atoms. In valence crystals this bond is much stronger.

Thus, it is only in molecular crystals that the bond is due to Van der Waals forces. In all other solids, the electrons are communalized in one form or another. True, in amorphous bodies we often find a superposition of bonds of diverse nature. Glass exhibits a simultaneous valence and ionic interaction. In complex organic compounds we have a combined valence and molecular interaction.

The end of the chain of questions

After what has been said about the forces between atoms and molecules and about the structure of solids, it is now a simple task to answer the question of why elastic forces arise when a table bends. (It is not easy to construct a quantitative theory, though; but we are not out to do that anyway).

Whether the table is wooden, plastic or metal, the atoms always come closer together upon compression, thus giving rise to a repulsive force. An elastic force originates. The forces of adhesion between atoms and molecules will inhibit tension.

A book on a table presses on the upper part of the wood and stretches the lower part. This continues until the distances between atoms are altered sufficiently for the elastic force to balance the book. If the book is removed, the interatomic distances are restored, and the surface of the table resumes its original shape.

Everything is very simple. When you grasp the main thing—the nature of electric forces between neutral systems—then you fully understand the origin of elastic forces.

Dry friction

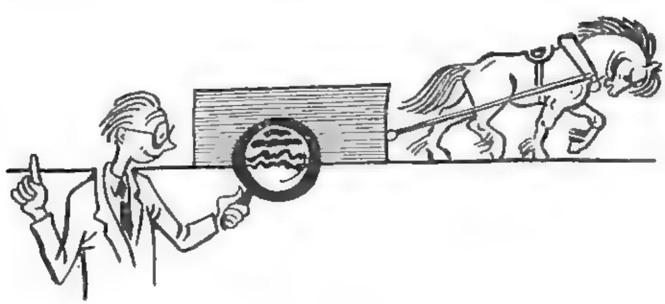
Now we know why a book does not fall through a table, but lies on top of it. But what keeps it from slipping off when the table is slightly inclined. Friction, of course.

At first glance, it is very easy to account for the force of friction, since the surfaces of the table and of the book are rough. This is apparent to the touch, and under a microscope the surface features are reminiscent of a mountainous country. The numerous projections interlock, become deformed and do not let the book slide. Thus, the static frictional force is due to the same forces of interaction of molecules as ordinary elasticity.

You increase the inclination of the table, and the book begins to slide. Obviously, the projections give way, molecular bonds are ruptured and are not able to withstand the increased load. The frictional force continues to operate, but it is now sliding friction. Broken projections are everywhere in evidence as the wear of rubbing surfaces.

It would seem that the more polished the surface, the smaller the force of friction. Which is true to a certain extent. Polishing reduces frictional forces between two steel bars. But there is a limit. The frictional force suddenly begins to build up as the surfaces become smoother. This is unexpected, but still can be explained.

As the surfaces become smoother, they fit closer together. However, so long as the heights of the projections exceed several molecular radii, there is no interaction between the molecules of the two surfaces, since these are short-range forces. Only for a high degree of polish do the surfaces come so close together that molecular adhesive forces become operative. They will inhibit the displacement



of the bars with respect to one another, thus ensuring static friction. When smooth bars slide over each other the molecular bonds between the surfaces are torn just as the bonds inside the projections in the case of rough surfaces. The rupture of molecular bonds is what distinguishes frictional forces from elastic forces, where there are no ruptures. That is why frictional forces depend on velocity.

Science-fiction stories often describe events in worlds without friction and we can see both the good and bad sides of this force. But don't forget that electric forces of interacting molecules lie at the heart of friction. Eliminating friction would actually involve dispensing with electrical forces, and, hence, the complete disintegration of matter.

Friction in liquids and gases

When two adjacent layers of liquid are in motion relative to one another, we have an ideal contact that is beyond the possibilities of solid surfaces, no matter how polished they may be. The molecules of the faster-moving layer entrain those of the slower layer due to molecular attraction, and the former are slowed down by the latter too. This is viscosity, or internal friction in liquids.

Due to the fluidity of liquids, not all molecular bonds are ruptured, as in the sliding of solid surfaces. Some of the molecules jump in the direction of action of molecular forces. The magnitude of friction will be in inverse proportion to the fluidity of the liquid and considerably less than static friction, as long as the relative speed of the liquids is not very great.*

In gases, the mean distance between molecules is so great that molecular attraction cannot give rise to friction between layers of a gas in relative motion. If the molecules did not fly out beyond the limits of these

* The physical processes both in dry and liquid friction are extremely complex, and no satisfactory quantitative theory of these phenomena exists as yet.

layers, there would be no friction. But thermal motion ejects the molecules outside the boundaries of the layers. Coming out of the fast layer into the slow layer, these molecules accelerate the latter in collisions, while slow-layer molecules that get into the fast layer decelerate it. Accelerations mean forces. But in gases the forces of friction are hundreds of times smaller than in liquids.

The forces of our muscles

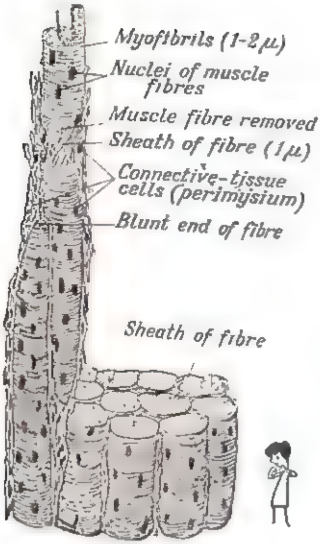
Gravity presses us to the earth, elastic forces hold us on the surface. Friction permits us to move about. Surface tension also helps. All of these forces are forces of nonliving nature, most of which we can control. Our control over them is due to the fact that we have at our disposal forces (of our muscles) that are subject only to reason and have no need of mechanical intermediaries.

The human muscle is one of the most remarkable mechanisms created by nature. First of all, it is an extremely efficient machine utilizing about 40% of the chemical energy it consumes. The best steam engine, by comparison, utilizes not more than 10%.

The force developed by a muscle deserves the greatest respect. Pick up a load and stand on one foot on tiptoe. The calf muscle is capable of lifting about 100 kilograms. Note that the foot can act as a lever and the muscle is attached to the short arm of this lever, so we now get nearly a ton. Also note that a human being cannot at will make it yield a maximum contraction. If its regulation by the nervous system is upset and the muscle is allowed to develop its full force, it will tear off a piece of the bone it is attached to.

Then, too, muscles have an enormous capacity for work. The heart muscle works day and night, without let up or any repairs for decades on end. This, so far, goes way beyond the capabilities of the best man-made machine.

Complicated chemical transformations inside cells lie at the heart of all muscle activity. We shall not go into all these problems, which are still largely unresolved, even despite the tremendous advances during recent years: four scientists have been awarded the Nobel Prize



for work on the chemistry of muscles. We shall confine ourselves to only one problem—where the muscle force comes from. What makes a muscle contract?

When you cut a steak, you can easily see that the muscle has a fibrous structure. Under the microscope one can see thousands of muscle fibres in the form of long cylinders laid in even rows. Each fibre is not one cell but a multitute of cells with a common cytoplasm and separated nuclei. The fibres are elongated strands—micelles—composed of bundles of protein molecules, which are the fundamental building material of living tissue.

A protein molecule is a long chain of alternating atoms of carbon and nitrogen. Two atoms of carbon and then a nitrogen atom, and so on. To the first carbon atom of every triplet is attached a more or less complex group of atoms. The entire molecule resembles a grapevine with leaves and tendrils extending in all directions, as the American physiologist Gerard in his book *The Body Functions* so aptly put it. We shall examine one of the many current hypotheses on muscle contraction. We cannot however give preference to any particular one.

The protein molecule is immersed in a liquid medium (the muscle is 70% water), whose chemical composition

determines whether the lateral groups of atoms link up forming temporary bonds or remain free. Due to the coupling of protein groups, the "spinal chord" of the molecule bends.

If there are a large number of such connections, the protein chain will bend and contract, like a grapevine when the tendrils and leaves are pulled together. This may be the mechanism that ensures muscle contraction.

The bonds between the branches and the trunk of the protein molecule are of a chemical nature. That is why the muscular force is essentially a chemical* force.

The lateral groups of atoms come together when the alkalinity of the solution surrounding the muscle fibre increases. The addition of some acid will make the muscle relax.

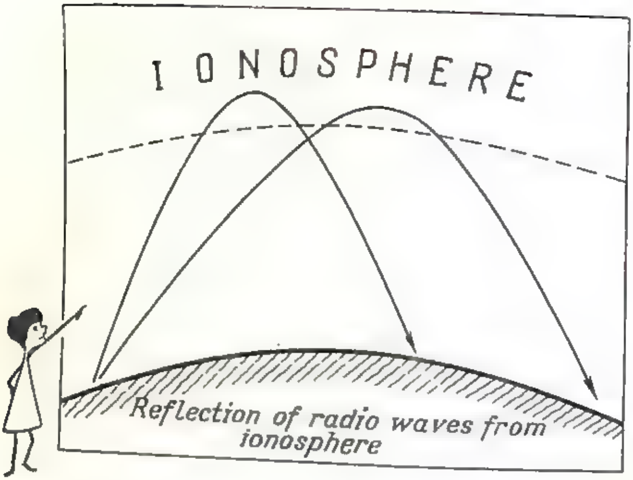
But how is it that from our desire to turn the page of this book a change takes place in the chemical composition of the medium in a large group of muscles? We do not have the answer. Neither does any other book, yet. The complicated series of processes that occur is still obscure in many stages.

4

Charged particles above us and around us

The natural state of bodies on the earth's surface (atoms and molecules and big pieces of matter too) is that of electric neutrality. However, if you charge an electroscope, it will soon lose the entire charge, no matter how perfect the insulation. Which implies that the air

* The chemical and hence electrical nature of muscle forces is the basis for other hypothetical mechanisms of contraction as well.



Reflection of radio waves from ionosphere

about us has a lot of charged particles, ions and pieces of dust. The ball of the electroscope "sucks up" ions of opposite sign and becomes neutral.

High up in the atmosphere is a thick layer of strongly ionized gas called the ionosphere. It starts some tens of kilometres above the earth and reaches to altitudes of four hundred kilometres. An electroscope is not enough to detect it. The discovery of the ionosphere was made by radio. This layer of highly ionized gas is a good conductor of electricity and, like a metal surface, reflects (short) radiowaves. Without the ionospheric mirror round the earth, radio communication via short waves would be possible only over distances of direct visibility.

Three suppliers

So we have ions up above us and all around us. But they are short-lived particles. An accidental encounter of two ions of different signs and they cease to exist. It is obvious, then, that there are certain constantly operating processes that supply us with ions.

There are three such suppliers. At the earth's surface, we have the emissions of radioactive elements found in the earth's crust in small quantities. At high altitudes there is the sun's ultraviolet radiation, and finally, the

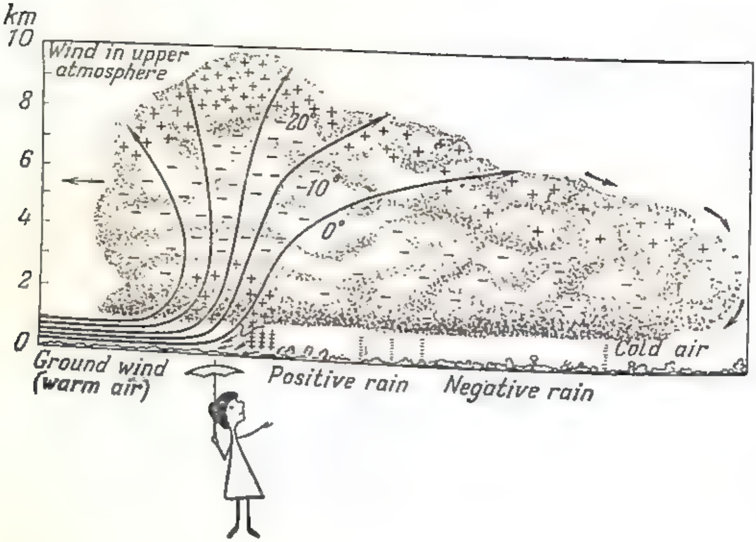
cosmic rays which pierce the entire atmospheric mantle from top to bottom. These are very fast charged particles. A small portion come from the sun and the rest from the deep cosmic space of our Galaxy.

Occasionally, our sun ejects powerful streams of charged particles. At several hundred kilometres above the earth, their electromagnetic fields excite atoms and make them emit light. We then see the northern lights, or auroras. Auroral displays are mostly seen in high latitudes. In temperate zones we hardly ever have the opportunity to enjoy the brilliant play of ribbons and columns of scintillating rainbow colours.

Lightning

But lightning is a familiar event. Enormous accumulations of electricity in a cloud give rise to a spark that extends at times tens of kilometres in length. Streaking along a bizarre pathway depending on the conductivity of the air and objects that it strikes, lightning produces weird effects. An assortment of marvellous lightning strokes is given by the French astronomer Flammarion in his book *The Atmosphere*.

"No theatrical effects or tricks can compare to lightning for sudden and weird displays," writes Flammarion. "It appears like a special kind of matter, something between the unconscious forces of nature and the conscious soul of human beings; it is a sort of spirit, delicate and fanciful, artful and dull at the same time, yet clairvoyant or blind, willful or constrained, hurtling from one extreme to the other, terrible and incomprehensible. It can't be talked to or caught. It *acts* and that is all. Its actions, of course, just like human actions, only appear to be capricious, actually however they obey strict laws. But we still haven't been able to catch hold of them. Lightning can kill outright and burn up a person, not only sparing but not even touching the clothes on his body. At another time it strips off the clothing without doing the slightest damage, inflicting not so much as a scratch. Then again it will steal all the coins from a purse without touching the pocket or the purse. It can tear off the gilding from a chandelier and transfer it to the walls of the room; or it takes the shoes off a traveller and



throws them ten metres to the side, or, finally, it drills through a pile of dishes, skipping every other two dishes... What kind of order can one establish here."

Then about a hundred cases are listed. For instance, "Lightning found one hairy man near E. and stripped the hair in bands the whole length of the body, then rolled it up into balls and tucked it deep into the muscles of the calves." "In the summer of 1865, a certain doctor in the locality of Vienna was returning home from the railway station. Getting out of the carriage he reached for his purse, and found that it had been stolen.

The purse was made of turtle skin and on one side was an incrusted steel monogram: two intertwined D's.

A short time later the doctor was called to examine a foreigner who had been "killed" by lightning and found unconscious under a tree. The first thing the doctor noticed on the patient's leg was his own monogram as if just photographed. Imagine his surprise. The patient was brought to and taken to the hospital. There the doctor said that his turtle purse ought to be in one of the patient's pockets, which turned out to be true. This man was the

thief who had stolen the purse, and electricity had stamped him by melting the metal monogram."*

In the United States an amazing case was recently reported. Lightning had hit an icebox and roasted the chicken inside, which was then cooled, because the refrigerator remained intact.

One may have some doubt about all these cases, but we can be sure about one thing—lightning is indeed capable of doing wonders, though it is not always possible to find explanations for the things it does. The discharge lasts only about a hundred thousandth of a second, and there is no way of getting prepared to observe it. There can be no repetition, because lightning never does exactly the same thing twice. Then, again, other circumstances change too.

Fundamentally, however, there is no great mystery at all. In the final analysis, everything reduces to such ordinary results of electric current as heating, an electromagnetic field, and chemical reactions. Only the current is enormous: tens and even hundreds of thousands of amperes.

The main thing is not to analyze all kinds of oddities in lightning's behaviour, but to grasp the general fact of how an electric charge builds up in a storm cloud. What causes electrification of water droplets and why are opposite charges spatially separated inside the cloud? Not everything is clear as yet.

First of all, there is no single mechanism for charging the droplets.

Several mechanisms have been definitely established and it is difficult to state which one plays the basic role. Here are two. A drop of water is polarized in the electric field of the earth (we have already mentioned that the globe is negatively charged). A positive charge accumulates in the lower part of the droplet and a negative charge in the upper portion. When a large drop falls it picks up mainly negative air ions and acquires an electric charge.

---

* A curious fact about Flammarion's statistics is that the number of women killed by lightning was only about one-third that of men. This of course has nothing to do with any preference on the part of the lightning; simply in those days (at the beginning of the twentieth century) more men worked in the fields of France than women.

Positive ions are carried upwards by ascending air currents.

Another mechanism is one in which droplets are charged as they are broken up by counter currents of air. The small spray is charged negatively and carried upwards, the larger particles are charged positively and fall downwards.

Both of these mechanisms provide for charging the water droplets and for separating charges of opposite sign within a cloud. Ordinarily, there is an accumulation of negative charge in the lower part of a storm cloud (with the exception of a small positively charged region) and an accumulation of positive charge in the upper portion.

Not so satisfactory are explanations of ball lightning, which occasionally appears after a strong discharge of linear lightning. It originates as a brilliant ball about 10 to 20 centimetres across. It often reminds one of a tiny kitten curled up and rolling along without feet. Ball lightning can explode when it comes in contact with objects and do considerable damage.

Ball lightning is probably the only gross phenomenon on earth that still lacks a satisfactory and reliable explanation. So far it has not been possible to produce ball lightning in the laboratory. That is the crux of the problem.

St. Elmo's fire

Just before a thunderstorm or during one, high pointed objects are seen to burst out as little cones of light. This slow and restrained discharge has since long ago been known as St. Elmo's fire.

The Roman historian Titus Livius has left us descriptions of this fire sparkling up on the masts of the galleys leaving port to fight the Athenians. The ancients believed the appearance of St. Elmo's fire to be a good omen.

Mountain climbers very often witness this fire as it spits out of metallic tips and even from the ends of the hair on their heads, embellishing them with tiny luminous headgear. If you raise your arm, you will have a burning sensation as the electricity streams off your finger tips. Pickanes often hum like bumblebees.

St. Elmo's fire is nothing more than a type of corona discharge that is simple to produce in the laboratory. A charged cloud induces electric charges of opposite sign on the earth's surface below it. Larger charges accumulate on pointed objects. When the electric field strength reaches the critical value of 30,000 V/cm, a discharge begins. The electrons which form about the top due to ordinary ionization of the air are accelerated by the field and break up the atoms and molecules they encounter. The number of electrons and ions builds up like an avalanche and the air begins to glow.

The electric charge of the earth

A storm cloud does not keep its charge for a long time. It is completely discharged after a few strokes of lightning. Disregarding slight deviations, the charge of the earth remains constant. At the earth's surface the electric field is not so small: 130 V/m. This is rather strange at first glance. Due to atmospheric ions the air conducts electricity, and calculations show that the whole globe should fully discharge in about a half hour. Therefore, the main difficulty is not to account for the origin of the charge, but to understand why it does not disappear.

There are two reasons why the earth's charge is restored. Firstly, lightning. Every day there are 40,000 thunderstorms ranging on the earth and every second 1,800 strokes of lightning hit the earth. The lower part of a cloud carries a negative charge and, consequently, a flash of lightning is simply a transfer of a portion of negative electricity to the earth.

At the same time, during a storm, electricity streams off numerous sharp-edged and pointed objects (St. Elmo's fire) which remove a positive charge from the earth's surface.

It is rather hard to strike a balance here, but on the whole things even up. The loss of negative charges by sections of the earth with a clear sky above is compensated for by an influx of negative charges in places where thunderstorms rage.

But where did the earth get its charge and why is it negative? We have to conjecture. Frenkel believes that at first a small charge was produced accidentally. It then began to build up via the lightning mechanism we have just described until the dynamic equilibrium set in that we still have.

The original charge might have been positive. And then the water droplets of a storm cloud would be polarized differently, and lightning strokes would deliver positive charges to the earth. It would be the same as we have now, only positive and negative charges would be interchanged.

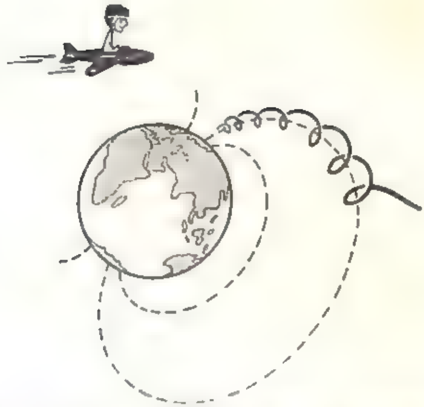
Terrestrial magnetism

The magnetic field of the earth was detected long before the electric field. It is very easy to detect, but its role is not at all exhausted in the help it offers to travellers in finding direction by compass at sea, in the taiga or in the desert.

While the electric field of the earth is confined largely within the lower layers of the atmosphere, the magnetic field extends out to 20-25 earth radii. Only at heights of about 100,000 kilometres does it cease to play a perceptible role and comes close to the magnitude of the field of interplanetary space.

The magnetic field forms a third protective belt about the earth in addition to the atmosphere and ionosphere. It helps to hold off streams of cosmic-ray particles when their energies are not too great. It is only in the vicinity of the magnetic poles that these particles can enter the atmosphere without opposition.

At large distances away the magnetic field is weak, but it embraces vast regions of space. It acts on a charged particle for a long time and is able to alter the trajectory quite substantially. Instead of a straight line we get a helix that spirals in along the lines of force of the field. The magnetic field drives the particles to the poles along these lines of force. At times, true, the velocity of the particle is very great and it does not even make a single circuit. Then we speak of the curvature of the trajectory.



In accord with Ampere's law, a magnetic field does not act on a particle flying along a line of force. For that reason, particles can freely move towards the poles where the lines of force spread out like a fan. No wonder then that corpuscular streams from the sun give rise to a glow in the upper layers of the ocean of air mainly at the poles.

Incidentally, these particle streams create their own considerable magnetic fields and generate magnetic storms that make compass needles run amok.

The radiation belts of the earth which were discovered just recently by space vehicles are nothing other than charged particles of not very large energies captured by the magnetic trap of the earth. The magnetic field holds these swarms of charged particles at great heights in the form of halos about the earth. Electrons predominate in the outer belt, and protons in the inner belt where the field strength is greater. These belts are dangerous for spacemen flying at high altitudes.

The earth as a spherical dynamo

The origin of terrestrial magnetism is still more obscure than that of the electric field. One cannot account for it as due to the accumulation of magnetized rocks. Frenkel recently advanced an interesting idea that appears to explain a great deal. The core of the earth is a generator

of electric current, operating on the principle of self-excitation, like an ordinary dynamo.

Let us recall this principle. In the dynamo, current is generated when conductors move in a magnetic field, which itself is set up by this very current. If there is no current at the start, it will appear when a certain rate of rotation has been reached and will continue to build up. The point is that there is always a small residual field. It creates a current which slightly increases the magnetic field. This then increases the current and then the magnetic field, and so forth until a limiting value is attained.

To liken the earth to a generator we must assume that the core is fluid and capable of conducting electricity. There is nothing unlikely in these presumptions. But where do the motions of the conducting masses of the core come from? In the dynamo we simply turn the armature, but in the core there are no outside agents.

Here is a way out. The temperature at the centre of the core should be somewhat higher than at the outer fringes due to radioactive decay of unstable elements. This gives rise to convection—the hotter masses rise upwards out of the centre and the cooler portions descend. But the earth is in rotation and the velocity of masses on the surface of the core is greater than deep inside. Therefore, the rising sections of the fluid inhibit the rotation of the outer layers of the core, and the descending portions accelerate the inner layers. As a result, the inner portion of the core rotates faster than the outer and plays the role of rotor (of a generator), while the outer part plays the role of stator.

Calculations demonstrate that this system is capable of self-excitation and the appearance of eddy electric currents of considerable magnitude. *According to Frenkel's hypothesis, these currents produce the magnetic field of the earth.* The energy required to maintain the current comes from the radioactive heating of matter that produces convection currents in the core.

All that may be true, but it is hard to say definitely. At any rate it is more correct to call the earth a large dynamo than a big magnet the way some books do.

The sun and other stars also have magnetic fields, which leave their imprint on the light waves they emit thus permitting physicists to detect them.

As Soviet spaceships have shown, the moon does not appear to have a magnetic field, at any rate it is at least 500 times weaker than the earth's field. No magnetic field has been found in the case of Venus or Mars either. So far we know nothing about the other planets of the solar system. We hope future space vehicles will give us this information.

The electrodynamics of space

Having taken up the topic of the magnetic fields of the planets and the stars, we enter a new realm, that of cosmic electrodynamics. There are as yet very few reliable facts, more hypotheses. But much of what just a little while ago was fanciful conjecture has now become almost reliable fact. The main thing is that electromagnetic forces play a much more important role in cosmic affairs than had hitherto been supposed.

The raging surface and atmosphere of the sun. Gigantic tongues of incandescent matter hurtling upwards. Whirls and tornadoes the size of our own planet. Incessant storms of fire and flame, storms of matter and of the magnetic field.

Occasionally black spots come out of the depths of the sun in pairs. In these spots, the magnetic field increases thousands of times over.

Sometimes unknown forces pull out of the sun whole blobs of charged particles and hurl them into the atmosphere of the earth against the gravitational field of the sun at velocities of several thousands of kilometres a second.

It is difficult to recognize law and order in situations like these. It is hard to grasp the nature of forces in whirling masses of matter. All this occurs very far away and under conditions that differ so greatly from those here on earth.

It is difficult though not impossible. At solar temperatures there are no neutral atoms, no neutral molecules. They simply cannot hold together, just like two trains colliding at full speed.

Now this totally ionized gas (or fully ionized plasma, as physicists term it) is an excellent conductor of elec-

tricity. This enables electromagnetic forces to act up and demonstrate their full strength in a new realm.

Considerable electric currents are generated in a magnetic field inside a moving high-temperature plasma. The conductivity is so good that these currents do not die out. And so in this medium, in addition to ordinary forces of elasticity, we have very important forces of the magnetic interaction of currents. Whereas motion in a simple medium can be described by the laws of hydrodynamics, here the dominating force is that of magnetohydrodynamics.

We are of course very far from any full understanding of events on the sun, but we are confident that the principal phenomena, ranging from the ejection of enormous masses of matter to the appearance of sunspots, are due to magnetic interactions.

But this is not all. Interstellar gas is highly ionized due to radiation. The density is low (one particle per cubic centimetre), but this is balanced by the enormous dimensions of the clouds. One has to take into account electric currents and hence magnetic fields here.

Moving clouds fill our entire Galaxy, and for this reason it is permeated by magnetic fields. This also holds for neighbouring regions of space as well.

The magnetic fields here are not large and we are not able to perceive them. But we know that they exist. How?

Radioemission of the Galaxy and cosmic rays

If we could see radiowaves, there would be three suns (actually radiosuns) in the sky instead of only one. One is located in the constellation of Cassiopeia, another on Cygnus and our own sun.* In addition, we would see a multitude of other less bright radiosuns and a faint scattered radiolight coming to us from all corners of the Galaxy and even from adjacent spaces that appear to be empty.

* The sun is an ordinary star; only the fact that it is so close enables it to compete in radio brightness with the other two sources which are immeasurably more powerful than our sun.

Some radiowaves are generated in collisions of charged particles of incandescent gas. This is thermal (or braking) radiation. It tells us nothing about the magnetic fields of the Galaxy. But there is another, nonthermal, portion that is due to the magnetic field, which spirals fast cosmic electrons. These electrons swing round the spiral and emit electromagnetic waves like a knife-sharpener puting on a display of sparks when a sharp knife edge touches the whirling stone. We may say that there are always magnetic fields to be found if radiowaves are generated.

But where do fast electrons come from in the cosmos? They give rise to radio emissions. Wherever we find powerful sources of radiowaves, we must seek cosmic accelerators. In other words, those distant and powerful radiosuns that we spoke of are, in the main, just such cosmic accelerators.

We are all accustomed to the deep and pure calm night sky. Nothing appears so permanent as the stars strewn over the firmament. This is really so. But occasionally explosions occur. And on a truly cosmic scale. A star that has for thousands of millions of years been living a routine existence suddenly and inexplicably bursts out to fantastic dimensions. (If this happened to our sun—though there is no danger, for ours is not of a type that will explode—it would expand beyond the boundaries of the outermost planets.) The brightness of such stars (they are called supernovae) increases hundreds of millions of times, making them visible even in full daylight. Gradually the brightness subsides and a nebulous cloud remains that is hardly distinguishable in a telescope.

In the Galaxy with its thousands of millions of stars, such an outburst occurs every 100 to 200 years. Since the telescope was invented there has not been a single new supernova.

So in most cases, radiosuns are the remnants of supernovae. What we observe in the direction of the constellation Cygnus, however, is a still greater catastrophe—the explosion of a whole Galaxy the size of our own.

We can imagine how the charged particles (electrons, protons, and atomic nuclei) are originally accelerated by a gigantic shockwave accompanying the outburst of the supernova. From then on, electromagnetic forces become operative. Building-up magnetic fields induce an electric

field. This field may not be so very large, but due to its tremendous cosmic dimensions it speeds up individual particles to energies far greater than those attained by the biggest man-made atom-smashing machines.

A certain portion of cosmic radiation is generated by the less powerful induction electric fields of the sun and other stars.

There is probably yet another mechanism of accelerating cosmic-ray particles. A moving magnetized cloud of interstellar gas encountering a fast-moving particle is similar to the collision of two spheres. Only here the role of ordinary elastic forces is played by the interaction of the particle with the electric field induced by the magnetic field moving with the gas. In such a collision, the energy of the particle should increase in the same way as a lightweight ball increases when struck by a very heavy one. By means of a large number of collisions the particle can acquire a substantial quantity of energy.

The random magnetic fields of the Galaxy not only accelerate but also scatter cosmic particles, making them come to earth uniformly from all directions and not only from their point of origin. Most likely, superhigh-power particles come from neighbouring Galaxies.

We are not positively sure that what we have described is exactly true. It is simply the most natural picture of electromagnetic phenomena in the universe today. It is a very rough picture, and not only because of its vast scale but also because the scientists who painted it are themselves not so sure of the details. Then again, the paint is hardly dry, so to say. The picture was painted just a short time ago, several years back, to be precise, and only its integrity gives hope that essentially it is correct.

The authors have a talk of their own

While momentous events on the cosmic scale proceed apace, the authors, working away in their Moscow flats, were not seeing eye to eye on a number of points. Contradictions arose and viewpoints began to diverge.

As will become clear from what follows, the essence of the argument boiled down to a clash between one of the authors (Mr. Acquiescent) and the other (Mr. Obstinate).

A. You know how I respect you, but this is no way to carry on. In place of a story about the essence of forces, you have a simple list of the various manifestations of electromagnetic forces. And then to this you add all sorts of oddities, many of which—you will forgive me—you yourself are not acquainted with. Surely the reader didn't buy the book for such matter. Do you think he needs just another textbook?

O. I'm sorry, but since the book was not approved by the Ministry of Education, it is not a textbook. What is more, didn't we promise a story about forces in nature? The forces that surround us? There is definitely no way of getting around friction, elasticity, the chemical forces, and so forth. Remember, we are not writing for philosophers interested only in fundamental principles, and not attracted at all by what is happening all around us all the time and every day.

A. I'm sure you have good intentions, but if you continue the way you're going, you'll have to discuss friction in liquids and the friction of solids, a ball, cylinder, cube, and the like. I may be exaggerating a little, but you definitely have that trend of mind—to pigeonhole every single thing.

O. What do you suggest, putting everything down to electricity, in just that one short word? Elasticity is due to electricity, friction is electricity, and chemical forces are really only electric forces. Is that what you want?

A. But look what you have. The structure of gases and the structure of liquids (which is familiar to everyone anyway), and the peculiarities of the forces in crystals

(which no one knows anything about and which hardly anyone is interested in)...

If you insist, of course, then go ahead, but write so that the reader doesn't fall asleep or throw the book out of the window.

O. Now listen, this is no easy job. It is of course much easier and more interesting to write about the theory of relativity than about chemical forces. Then each type of electromagnetic force requires a whole book in itself. Of course, things get dull if the story is too brief.

A. It is not only more interesting to write about the theory of relativity, it is more interesting to read about it, as well.

What upsets me is that you are perfectly calm about being dull. Why? After all, there are any number of encyclopedias where one can read about anything under the sun.

O. Good and well, let this part of the book be written dictionary-style. Still it shouldn't be too hard to get through.

A. You really are obstinate, I see. The trouble with you, though, is that you don't follow any logical sequence of ideas. After cosmic radiation you want to go over to electric fish.

O. So what? Fish are interesting to read about. If the reader doesn't think so, he can skip that section. Perhaps we ought to write in the introduction that the reader can just choose what he wants in this chapter, or he can even skip the whole chapter if he wants to.

A. If he does that, he won't read our interesting conversation—too bad.

Electric fish

Electric fish are unique creatures that differ from their brethren in that they carry living galvanic cells with them. The electricity they generate serves as a form of defense and attack.

It is interesting to note that among fossil fishes there were more electric species than presently exist. Apparently, the use of electromagnetic forces was not so effective as perfecting other systems that are not so obvious, first of all, the muscular systems.

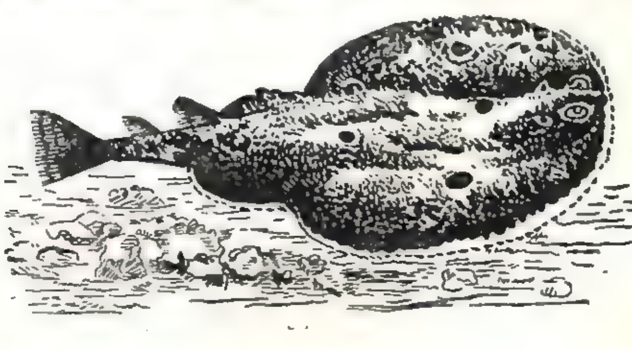
The most brilliant representative of this kind is the torpedo fish, which lives in warm seas, weighs about 100 kilograms and grows to about two metres in length. Its electrical organs, which are located laterally along the head, weigh over 16 pounds. A torpedo fish is able to produce an electric shock of 8 amperes at a voltage of 300 volts. This is definitely dangerous to human beings.

The natural thing is to suppose that electric fish are not very sensitive to electricity. Which is actually the case, for the torpedo fish can easily stand up to voltages that kill other fishes.

The electric organs of the torpedo fish are remarkably similar to a battery of galvanic cells. They consist of numerous plates arranged in piles (the cells are connected in series) which are situated one after the other in a number of rows (this is a parallel connection). One side of each plate is smooth and carries a negative charge, the other side has projections and is charged positively. The whole set-up is contained in an insulating tissue, as one would suppose.

We shall not try to delve deep into the mechanism for developing electromotive force in the organs of the torpedo fish, in the same way that we dispensed with analyzing the operating principle of the ordinary galvanic cell. There are many obscure points here. We are sure about one thing, however, and that is that (like in the galvanic cell) chemical forces lie at the heart of the operation of these electric organs.

We shall not take up any more electric fishes, but just must dwell for a moment on a remarkable inhabitant of the Nile River—the mormirus. This fish is equipped with a marvellous radar system. At the base of the tail



it has a generator of alternating electric current that sends out impulses with a frequency of several hundred oscillations per second. The surrounding objects distort the electromagnetic field about the fish, and this is immediately recorded by a receiving device on the back. This radar system is extremely sensitive. The mormirus cannot be caught in a net. In an aquarium, it will become agitated even when you draw a comb through your hair.

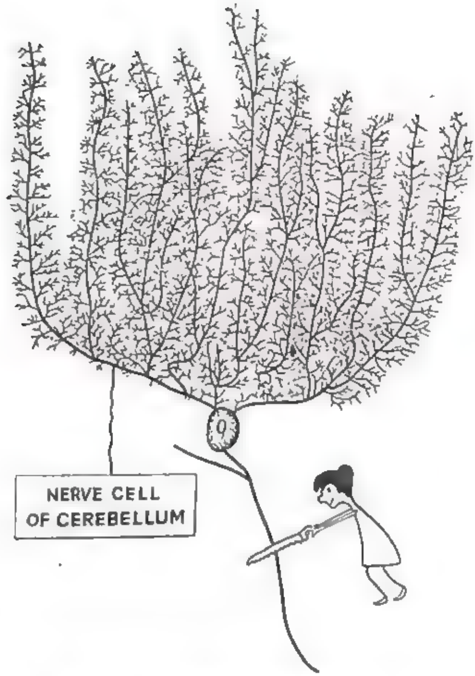
We still don't know exactly how the radar works, but there is hope that a detailed study of the problem will help perfect submarine electromagnetic communications, which has long been a stumbling block due to the rapid attenuation of electromagnetic waves in water.

The nature of a nerve impulse

The torpedo fish and other electric fishes are really nothing more than a freak of nature. Free electricity in living organisms plays a much more important role. This is the electricity which maintains communication lines that transmit telegrams to the brain from the sense organs about what is going on in the surrounding world, and the responses of the brain in the form of commands to the muscles and the internal organs.

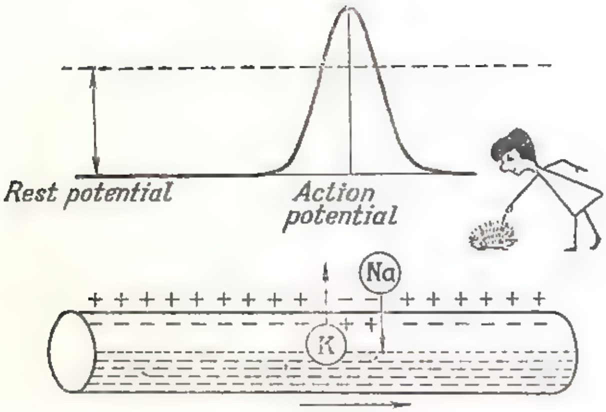
Nerves permeate the whole body of all higher living beings and make the organism into an integral unit that operates with marvellous expediency. When a nerve leading to some muscle is out, the muscle is just as paralyzed as the cylinder of a motor when the wire leading to the spark plug is severed.

This is not simply an external analogy. Since the time of Galvani it has been known that a nerve pulse carried over nerve fibres is simply a momentary electric impulse. True, things are a bit more complicated than might be thought. The nerve is not a passive canal of high conductivity like an ordinary metallic wire. It is actually more like a relay line, where a signal is transmitted to the next link, where it is amplified and then sent onto the next link, where it is again amplified, and so on. In this way, the signal can be delivered to considerable distances in full strength, despite natural attenuation.



What is a nerve? Here is Gerard's description: "If a spider, hanging just above the ground by its five thread attached to the top of a six-story building, were reduced in size about twenty times, thread included, it would rather resemble a nerve cell, or neurone. The main body of the nerve cell, with its nucleus and cytoplasm, is not unduly large or peculiar as cells go... But the neurone does not stop with its cell body, as ordinary uninquisitive cells do, but sends out five thread-like processes to explore far distant regions. Most of the processes extend for moderate distances... But one slender process, often less than 0.01 millimeter in diameter, suffers from unbridded wanderlust and threads its course for centimeters or even meters away from its origin.

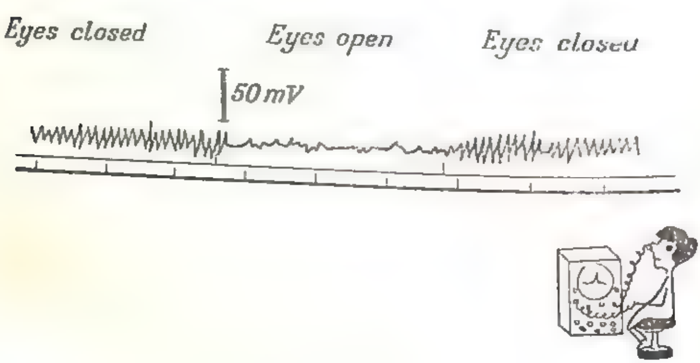
All the neurones of the regular nervous system are gathered together in the central brain and spiral chord, of which they form the gray matter... Only their long processes, the axones (from axis), connect them with the rest of the body. Bundles of these axis cylinders, coming from related cells, make up the nerves."



A special substance called myelin envelopes most axons in a thin layer, like insulation tape around an electric wire.

The axon itself can be pictured roughly as a long cylindrical tube with a surface membrane separating two aqueous solutions of different chemical composition and concentration. The membrane is like a wall with a large number of doors ajar through which ions of the solutions can move with great difficulty. The most remarkable thing (and at the same time the least understood) is that the electric field closes up the doors, but as the field weakens the doors open wide.

There is an excess of potassium ions inside the axon when in a state of rest; on the outside is an excess of sodium ions. The negative ions are concentrated mainly



on the inner surface of the membrane and for that reason it is charged negatively, while the outer surface has a positive charge.

When a nerve is stimulated, the membrane is partially depolarized (a reduction in charge on its surfaces), which brings about a fall in the electric field inside it. The doors thus open for the sodium ions which begin to enter the fibre. Finally the inner part of the axon becomes charged positively on this section.

That is how a nerve pulse develops. Actually, it is a pulse of voltage caused by the passage of current through the membrane.

At this instant, the doors open for the potassium ions. They pass onto the surface of the axon and gradually restore the voltage (about 0.05 volt) that the nerve had in the unexcited state.

At the same time, part of the ions from the neighbouring section break through the adjacent doors, weakening the field here. And the entire process repeats itself on a new section of the axon. In this way, a nerve impulse moves towards the brain without attenuation at a velocity of about 120 metres per second.

The ions of sodium and potassium, which were displaced from their regular sites, gradually return through the wall due to chemical processes that are still rather obscure.

The amazing thing is that the behaviour of higher animals, all the creative efforts of the human brain are based, in the final analysis, on these exceedingly weak currents and extremely subtle microscopic chemical reactions.

Brain waves

We now turn to the sanctum sanctorum of living nature—the human brain. Electrical processes are constantly active in the brain. If metallic plates are applied to the forehead and back of the head and connected through an amplifier to a recording instrument, we can register continuous electric oscillations of the cerebral cortex (this goes for animals as well as humans). Their rhythm, shape and intensity depend very significantly on the state of the person.

The brain of a person sitting completely at rest with eyes closed and thinking of nothing will register about 10 oscillations per second. When the person opens his eyes, the brain waves disappear, only to reappear again when the eyes are closed. When a person falls asleep, the rhythm of oscillations becomes slower. The type of wave gives a very precise indication of the onset and end of dreams.

In the case of brain diseases the character of the electric waves undergoes a marked change. Thus, pathological oscillations in the case of epilepsy can serve as a definite indication of the disease.

From the foregoing it is obvious that the brain cells are in a state of constant activity, and large quantities of them vibrate together like the violins of an enormous orchestra. The nerve impulses entering the brain do not proceed by familiar pathways, but alter the entire pattern of wave distribution in the cerebral cortex.

The nature of electrical activity of the brain changes with age throughout one's life and learning.

We must presume that these electric oscillations do not simply accompany the work of the brain like the noise of a moving automobile, but represent a very fundamental factor of life as such. Electromagnetic processes lie at the heart of all processes in the electronic computer, which is capable of performing certain functions of the human brain better than the brain itself. It must be stressed, however, that there is no direct correlation between a specific brain wave and a difinite thought or sensation. There is no way of telling what a person is thinking about from the shape of the brain waves. *

So far we do not know what functions these processes perform in the brain. But they clearly demonstrate that electromagnetic processes in the most highly organized matter of nature serve as the material basis of thinking.

---

* Incidentally, for this reason and also because of the extreme weakness of electric oscillations, it is hard to believe that electromagnetic waves generated by these electric oscillations could be the basis of telepathy, about which we have been hearing so much of late.

5

The rays of the sun

"So dear to me are the sticky unfolding buds of spring, so dear the blue sky above," said Ivan Karamazov, one of the gloomiest characters created by the genius of Dostoyevsky.

The light of the sun has always been a symbol of eternal youth, of the best things of life. And what excitement and happiness are in the words of the little boy out in the sun:

Let there always be sunshine,
Let there always be sky,
Let there always be Mama,
Let there always be me!

or take this stanza by the poet Dmitri Kedrin:

You say our fire is dying out,
You insist we're old, you and I,
But look at the sparkling deep blue sky,
It's much much older than you or I!

The kingdom of darkness, the world of gloom—it is not simply a lack of light, it is a symbol of everything heavy and oppressive to the spirit of man.

Sun worship is by far the oldest and most beautiful cult of humanity. The magical god Kon-Tiki of the dwellers of Peru, the god Ra of the ancient Egyptians. At the very dawn of humanity, people realized that the sun means life. We have long since learned that the sun is not a god but an enormous ball of fire, yet man's awe will go on to eternity.

Even a physicist who is used to the precise recording of events experiences an unpleasant feeling of irreverence if one says that sunlight is nothing more than electromagnetic waves of a definite length. But actually that is so, and here in this book we shall try to say just that.

We perceive light, as electromagnetic waves of length 0.00004 cm to 0.000072 cm. The other wavelengths do not produce visual impressions.

Wavelengths of light are very short. Imagine a wave on the sea magnified so that it occupies the entire Atlantic Ocean from New York to Lisbon. If magnified to the same scale, a light wave would be just a bit longer than the width of this page.

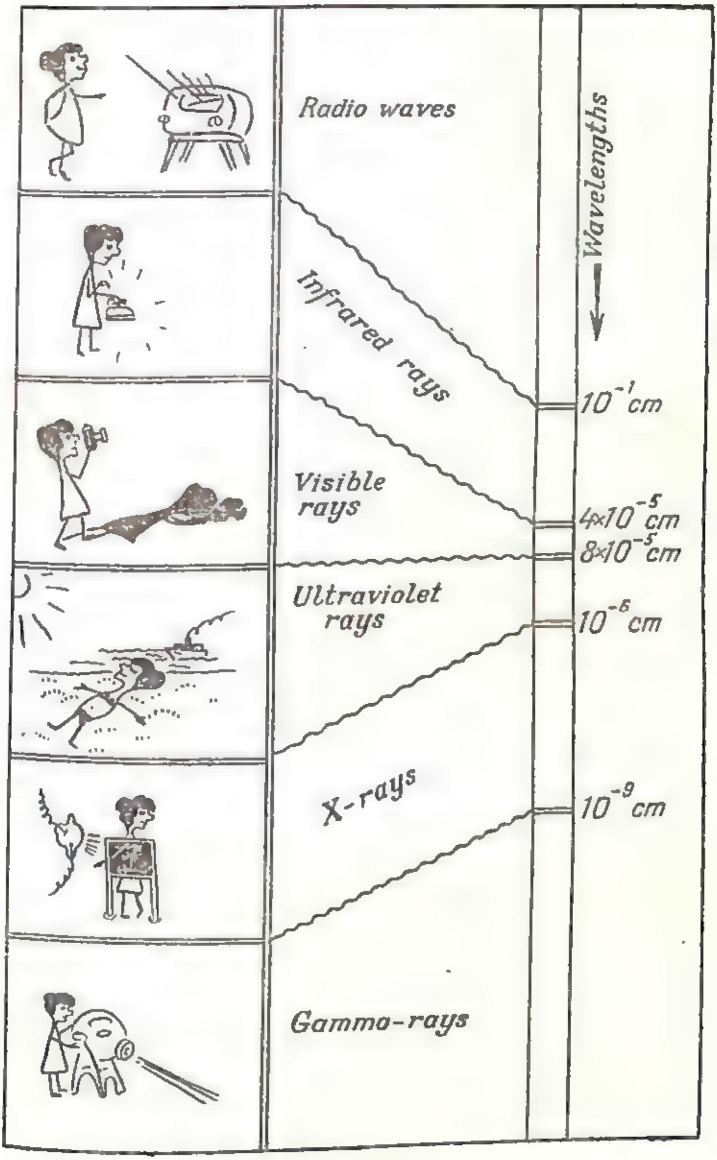
The eye and electromagnetic waves

We are naturally familiar with electromagnetic waves of quite different lengths. There are kilometre waves and there are waves shorter than visible light: ultraviolet rays, X-rays, and others. Why did nature make our eye (and the eyes of other animals) responsive only to a specific and very narrow band of wavelengths?

In the scale of electromagnetic waves, visible light occupies a tiny strip squeezed in between ultraviolet and infrared rays. Off both sides extend broad bands of radiowaves and gamma-rays emitted by atomic nuclei.

All these waves carry energy and, it would seem, are capable of doing for us the very same things that light does. The eye could be sensitive to them.

We can say right off that not all wavelengths are suitable. The gamma- and X-rays are emitted only under very special conditions and are hardly ever found



Radio waves
Infrared rays
Visible rays
Ultraviolet rays
X-rays
Gamma-rays
Wavelengths
10^-1 cm
4×10^-5 cm
8×10^-5 cm
10^-6 cm
10^-9 cm

around us. Which is very good, of course. They (especially gamma-rays) cause radiation sickness, so that man would not be able to enjoy the world for a long time in gamma-rays.

The long radiowaves would be very inconvenient. They easily pass round objects a metre or so in size, like the ocean waves skirting big rocks on the seashore, and we wouldn't be able to see many things that we find very necessary. This habit of waves to move round obstacles (diffraction) would result in our perceiving the world like a fish in the mud.

But then there are the infrared (thermal) rays that heat bodies though they are invisible. They might do the job of the visible wavelengths, or, finally, the ultraviolet rays might be harnessed.

So is this narrow band of wavelengths that we call visible light accidental? Particularly, since the sun emits both visible light and ultraviolet and infrared rays.

By no means! First of all, the maximum solar radiation of electromagnetic waves lies precisely in the yellow-green region of the visible spectrum. But even this is not the most important thing. The neighbouring regions of the spectrum would be strong enough.

Windows in the atmosphere

We live at the bottom of an ocean of air. The earth is surrounded by an atmosphere. We consider it to be transparent or nearly so. Which is actually true, but only for a narrow band of wavelengths (a small region of the spectrum, physicists say), just that portion which our eye can perceive.

This is the first (optical) window in the atmosphere. Oxygen intensely absorbs ultraviolet rays. Water vapour stops infrared radiation. Radiowaves are reflected from the ionosphere.

There is only one radio-window open to waves from 0.25 cm to about 30 metres long. But they do not affect the eye and, what is more, their intensity in the solar spectrum is very low. A big advance in radio engineering

was required (the radar of World War II) before these waves could be put to use.

Thus, in the process of evolution, living organisms acquired an organ that responds to the strongest radiations and does its job very well.

The fact that the greatest amount of solar radiation passes right in the middle of the optical window is surely a wonderful gift of nature. (On the whole, nature has been exceedingly generous as regards this planet of ours, doing everything possible to make life liveable. It could not, of course, foresee all the consequences of its generosity, but it gave us reason, the power to think, thus making us responsible for ourselves and our fate.) Possibly, we could have got along without the amazing coincidence of maximum solar radiation and maximum transparency of the atmosphere. The sun's rays would have started life on earth sooner or later and they would have kept it going too.

The blue sky

We assume the reader is reading carefully and we believe he may have recognized a contradiction in the fact that maximum solar radiation lies in the yellow-green region of the spectrum, though we see the sun as a yellow ball.

The fault lies with the atmosphere, which lets through more of the long-wave portion of the spectrum (yellow) and less of the short waves. That is why the green is greatly attenuated.

Generally speaking, the short waves are largely scattered by the air. That explains why the sky is blue and not yellow or red. If there weren't any atmosphere, there would be no sky as we know it. Instead there would be a black void and a brilliant sun. So far, only spacemen have had that view.

Our sun is dangerous if one is not protected. High up in the mountains the sun burns unbearably (the ultraviolet radiation is not absorbed sufficiently by the upper layers of the atmosphere). Without the protection of clothing and dark glasses, the skin and the retina of the eyes could easily be burnt.

The gifts of the sun

Light waves falling to earth are nature's great gift. They bring warmth and life. Without them the cosmic cold would freeze everything. If all the energy consumed by human beings (fuel, falling water and the wind) were increased 30 times over, that would make only one thousandth of the total energy that the sun supplies free of charge and without any effort on our part.

What is more, the main fuels (coal and oil) are simply stored solar rays, anyway. They consist of the remnants of a lush vegetation that once covered the earth and, partly, of the remains of the animal kingdom.

The water that turns the turbines of power stations was once raised as water vapour into the atmosphere by the energy of the sun's rays. It is again solar radiation that sets in motion the air masses of our planet.

But this is not all. Light waves do not only heat. They stimulate chemical activity that simple heating cannot bring about. Tan and the fading of fabrics are due to chemical reactions.

The most important of these reactions are those that occur in the "sticky buds of spring" and in the pine needles, blades of grass, the trees and in multitudes of microorganisms. Processes occur in the green leaf in the sun that make life on earth possible. They furnish food and oxygen.

Our body, like the organisms of other higher animals, is not capable of combining pure chemical elements into complex chains of atoms, the molecules of organic substances. In breathing we constantly pollute the atmosphere. We take in vitally important oxygen and exhale carbon dioxide. We thus fix oxygen and make the air unsuitable for breathing. The air must constantly be purified. This is done by plants on the ground and microorganisms in the seas.

Leaves absorb carbon dioxide from the air and break up its molecules into their component parts: carbon and oxygen. Carbon goes to build the living tissues of plants, while pure oxygen is returned to the air. Linking up to the carbon chain the atoms of other elements extracted from the earth by the roots, plants construct molecules of protein, fats and carbohydrates, which all serve as food for human beings and for animals.

All this is done at the expense of the energy of solar radiation. The important thing here is not only the energy itself, but the form in which it is supplied. Photosynthesis (that is what scientists call this process) can occur only by means of electromagnetic waves of a very definite interval of the spectrum.

We shall not attempt to describe the mechanism of photosynthesis. It still hasn't been deciphered completely. When it is, mankind will probably enter a new era, for proteins and other organic substances will then be grown in retorts under the blue sky.

Light pressure

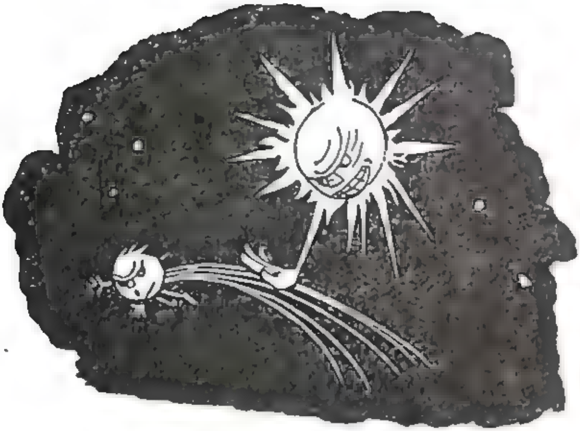
Subtle chemical reactions give rise to light. Light is capable of simple mechanical actions. It can exert pressure on bodies. True, light acts in its usual delicate way—the pressure is very small. Only about half a milligram per square metre of the earth's surface on a bright sunny day.

That comes out to only a slight force of 60,000 tons over the entire earth, which is practically negligible when compared with the gravitational force, which is 10^{14} times greater.

It required the talent of P. Lebedev to discover the pressure of light. This he did at the beginning of the century when he measured the pressure on solids and on gases.

Despite the fact that light pressure is very small, its action is sometimes apparent even to the unaided eye. You have to watch a comet to see what light pressure can do. The tail of a comet is made up of minute particles and is always seen to be pointing away from the sun. The particles of the cometary tail are so small that the forces of light pressure become comparable to or even greater than the attractive forces of the sun. That is why the tails of comets are always pushed away from the sun.

This is easy to account for. The force of gravitation is proportional to the mass and, consequently, to the cube of linear dimensions of the body. Now solar pressure is proportional to the surface area, hence to the square of the linear dimensions. As particles get smaller, the



gravitational forces decrease *faster* than the pressure, and for sufficiently small particles they become less than the forces of light pressure.

An interesting thing happened to the American satellite "Echo". When it settled into orbit, compressed gas filled a large polyethylene bag producing a lightweight sphere of about 30 metres in diameter. Unexpectedly, it was found that during one circuit about the earth the pressure of solar radiation displaced it five metres from its original orbit. As a result, instead of the planned 20 years lifetime, it stayed in orbit less than one year.

In stellar interiors where the temperature rises to several million degrees, the pressure of electromagnetic waves should reach fantastic magnitudes. Along with the gravitational force and ordinary pressure, it apparently plays a significant role in intrastellar processes.

The mechanism that generates light pressure is relatively simple and we can say a few words about it. The electric field of an electromagnetic wave falling on a substance puts the electrons into oscillation. They begin to vibrate transverse to the direction of propagation of the wave. This in itself does not yet give rise to pressure. The moving electrons are now acted upon by the magnetic field of the wave. It is this field that pushes the electrons along the light ray and that ultimately generates pressure on the piece of matter as a whole.

Heralds of distant worlds

In the vast boundless expanses of the universe, our Galaxy is just another assemblage of stars, and our sun is a typical yellow dwarf. Only within the solar system do we find a privileged planet—the earth. Privileged because it is most suitable to life.

We not only know the positions o numerous stellar worlds, we know what they are made up of too. They consist of the very same atoms as our earth. Ultimately, the world is everywhere the same.

Light is the herald of distant worlds. It is the source of life and the source of our knowledge about the universe. If the electromagnetic waves that come to earth could speak they would say: "How magnificent and beautiful the world is." Only electromagnetic waves bring in important information, the gravitational fields hardly tell us anything about the universe.

Stars and stellar clusters are visible to the unaided eye and are also observed telescopically. But how do we know what they consist of? This information we obtain from a spectroscope, which is an apparatus that sorts out the light waves dispersing them in different directions.

Hot solids or liquid bodies emit a continuous spectrum, which means that it includes all wavelengths, from the long infrared to the short ultraviolet.

Quite a different picture emerges when we deal with isolated or nearly isolated atoms of a substance in the form of incandescent vapours. The spectrum then consists of a series of coloured lines of different brightness separated by broad dark bands. Each coloured line is associated with some electromagnetic wave of a definite length (remember there are no colours in nature about us, only waves of different length).

The most important thing to bear in mind is that the atoms of a given chemical element yield a specific spectrum that is different from the spectra of the atoms of other elements. Like the finger-prints of human beings, the line spectra of atoms are uniquely individual. The unique pattern on the finger of a person helps to locate a criminal. In exactly the same way, the individuality of the spectrum enables physicists to determine the

chemical composition of a body without touching it, and not only when it is near at hand, but when it is millions of light years distant from us. The only requirement is that the body must shine brightly.*

The elements that we know on the earth have also been found on the sun and the stars. Helium was even detected in the sun before it was found here on the earth.

If the emitting atoms are located in a magnetic field, their spectrum is substantially altered. The separate colour bands are broken up into a number of lines. This is what enables us to discover a magnetic field in stars and to evaluate its magnitude.

The stars are so far away that we are not able to distinguish any motions they may have. But light waves coming from them carry this information as well. The dependence of the wavelength on the rate of motion of the source (the Doppler effect that we have already mentioned) permits us to evaluate stellar velocities and to determine any rotation they may have.

Most of the information we have about the universe comes to us through the optical window in the atmosphere. With the development of radio astronomy, more and more information about the Galaxy is coming through the radio window.

Where electromagnetic waves come from

We know (that is, we think we know) how radiowaves are generated in the universe. One of the radiation sources has already been mentioned: thermal radiation that originates when charged particles are decelerated in collisions. Of greater interest is nonthermal radio emission.

---

* The chemical composition of the sun and stars is determined not by the emission spectrum, since this is a continuous spectrum of the dense photosphere, but by the absorption spectra of the solar atmosphere. Vapours absorb most intensely those wavelengths which they emit in the incandescent state. The dark absorption lines on the background of the continuous spectrum enable us to determine the composition of astronomic bodies.

Visible light, infrared and ultraviolet radiation are all nearly exclusively of thermal origin. The high temperatures of the sun and other stars are what gives birth to electromagnetic waves. Stars also emit radiowaves and X-rays but their intensity is very low.

Collisions of the charged particles of cosmic radiation with the atoms of the terrestrial atmosphere produce short-wave radiation: gamma-rays and X-rays. True, though born in the upper layers of the atmosphere they are mostly absorbed during their passage downward and never reach the earth's surface.

The main source of gamma-rays at the earth's surface is the radioactive disintegration of atomic nuclei. Here, the energy is extracted from the richest energy store of nature—the atomic nucleus.

Electromagnetic waves are emitted by all living creatures as well. First of all, infrared rays, which are emitted by every hot body. Some insects (for instance, glow-worms) and deep-water fishes emit visible light. Here the light is generated in chemical reactions in luminescent organs (cold light).

Finally, chemical reactions associated with the division of cells of plant and animal tissues are the source of ultraviolet radiation. These are the so-called metogenetic rays discovered by the Soviet scientist Gurvich. At one time they appeared to be important in the vital processes of cells, but subsequent precise experiments did not corroborate these suppositions.

The sense of smell and electromagnetic waves

The sense organs are not affected solely by visible light. One can feel warmth at a distance from a hot stove or a teakettle. The human organism is capable of perceiving sufficiently intense fluxes of infrared rays. True, the sensitive elements in the skin do not respond directly to the radiation, but to the heating effects that it produces. Perhaps this is the only effect that infrared rays produce on the body, yet it may not be so either. The final answer will be forthcoming when we have solved the mystery of the sense of smell.

In what way does man, and particularly animals and insects, perceive the presence of certain substances at a distance by their odour? The obviously simple answer is that molecules of the substance stimulate these organs in a specific way producing what we know as odours.

How is this to be explained? Bees converge on honey even when it is hermetically sealed in a glass jar. Or take this fact: some insects perceive the odour of a substance of such small concentrations that there is less than one molecule per insect.

In this connection a hypothesis was advanced explaining the sense of smell as being due to electromagnetic waves that are over 10 times the length of visible light waves. These waves are emitted in low-frequency oscillations of molecules and affect the olfactory organs. A curious thing about this theory is that it unexpectedly brings together the eye and the nose. Both are types of receptors and analyzers of electromagnetic waves. However, we are still not sure whether this is so.

A cloud in the sky

By now the reader, who has learned about the almost infinite diversity of manifestations of electromagnetism, even going so far as the sense of smell, is probably sure that there is no more harmoniously constructed theory alive today. True, there was a little hitch in our discussion of the structure of atoms. Otherwise, electrodynamics appears to be flawless and unassailable.

Such was the feeling of perfect complacency at the end of last century, when physicists did not know anything about atomic structure. The general feeling was so overwhelming that the famous English physicist Thomson, at the turn of the century, spoke about a clear scientific horizon where he could discern only two tiny clouds. He had in mind the experiments of Michelson in measuring the velocity of light and the problem of thermal radiation. Michelson's experiments served as a basis for the theory of relativity. Thermal radiation will have to wait a bit.

Physicists were not surprised by the fact that hot bodies emit electromagnetic waves. The only thing was to

learn how to describe them quantitatively on the basis of the elegant system of Maxwell's equations and the Newtonian laws of mechanics. In the solution of this problem, Rayleigh and Jeans obtained a remarkable and paradoxical result. From the theory, it inevitably followed that even the human body at 36.6°C should be shining brightly and losing energy so fast that it would rapidly cool off to almost absolute zero.

There were no subtle experiments here to demonstrate the obvious conflict between theory and actuality. Yet the calculations of Rayleigh and Jeans left no doubts. They were immediate corollaries of the most general statements of the theory. Not even extremely involved manoeuvres could save the situation.

The fact that the laws of electromagnetism, which had been verified numberless times, balked as soon as they were applied to the problem of radiation of short electromagnetic waves was so amazing that physicists began to speak of an "ultraviolet catastrophe".* That was what Thomson had in view when he spoke of "clouds". But why only a little cloud? Well, because at that time physicists considered the problem of thermal radiation

---

* The catastrophe was called ultraviolet because the unpleasantness came to light in the study of very short waves.

a small particular problem of no great importance on the background of general magnificent achievements.

However, this "cloudlet" was fated to build up into an enormous storm cloud covering the entire scientific horizon, and expending itself in a terrific downpour that scoured away the entire foundation of classical physics. But at the same time it gave birth to a new physical conception of the world which we now designate as the quantum theory.

Before taking up this new theory which wrought such a revolution in our understanding of electromagnetic forces, and forces in general, let us glance back (from the height that we have now attained) at the road we have covered and try to picture why electromagnetic forces play such an outstanding part in nature.

6

Why are the two biggest chapters of this book one of which hasn't even got a number concerned with electromagnetic forces?

Obviously because electromagnetic forces are the most common in nature. The brilliant and possibly rather confused kaleidoscope of events in this chapter is sure proof.

What is the reason for this great diversity of electromagnetic manifestations? Why has nature provided these forces with such a vast stage? The answer to the second question is to some extent contained in the statement of the first. The multiplicity of forms of electromagnetic interactions of course contributes to participation in diverse processes of living and nonliving nature.

We are not going to say anything new. On the basis of what you have read you can answer the questions

yourself. Put the book aside for a moment and ponder a bit on the reason for such a diversity of electromagnetic forces.

Now see whether you have taken everything into consideration.

The existence of two types of charge (positive and negative) is obviously one of the most important factors in this multiplicity of forces. It makes possible both attraction and repulsion. If some positive charge is equal to a negative charge, then the bodies do not exhibit interaction over appreciable distances. Electromagnetic forces, though by nature long-range forces, can be also short-range.

Another factor that we must bear in mind is the relative complexity of the laws of electromagnetic interaction.

Unlike gravitational forces, *electromagnetic forces depend not only on the distance between charges, but also on the velocities.* There is a specific magnetic interaction that has no analogy in the Newtonian theory of gravitation.

Finally, a charged body undergoing acceleration generates electromagnetic waves. *Interaction depends on accelerations.*

However, diversity of form would not count for much if all bodies were not composed of electrically charged particles. *The component parts of the atom—the nucleus and electrons—carry electric charges.*

All particles without exception are endowed with a gravitational charge (mass), but the forces of gravitation are exceedingly weak and offer no competition whatsoever to the powerful electromagnetic forces within pieces of matter.

Stronger still are nuclear interactions. But they operate only over very small distances. Even between neutral systems electromagnetic forces greatly exceed nuclear forces as to range of action, while forces between charged bodies are no less long-range than are the forces of gravitation. And interaction due to electromagnetic waves falls off with distance still more slowly.

These reasons are quite sufficient to make electromagnetic forces the most "popular" forces of nature.

7

Actually, this should be a whole chapter, so important are the forces we shall discuss. They have to do with the laws that govern the world of elementary particles, which make up all the objects and things around us. The laws of interaction of these particles ultimately determine the "forces of nature" that make up our story. There are many reasons why we did not begin with this section and did not even turn it into a separate chapter. First of all, the pathway from the complex to the simple is not always the best; the study of mathematics begins with arithmetic and not with integrals; why scare the reader from the very start; and so on and so forth, and, finally, there is no reason to be overpedantic.

There was yet another rather important consideration. One finds it possible to appraise a physical idea in full measure only when the inner logic of its appearance is grasped and its place in the general cognitive chain of the laws of nature is found.

Now that we have completed the story of electromagnetic forces in their classical interpretation, let us open another door and peer into a truly fantastic world, the world of the ultrasmall, or the *microworld*, as it is called.

Discontinuity in the continuous

Science has its own symbolism. The word "quantum" came in with the twentieth century, and anyone interested in the biography of ideas will see how exciting, even tragic, the history of this term is.

Max Planck was already a mature scientist when he got interested in the problem of the radiation of electromagnetic waves by hot bodies. Planck's education, like that of other scientists of his generation, was based entirely on what appeared to be the harmonious and complet-

ed picture of the world called classical physics. The firm foundation here was the Newtonian concept of motion, and even the rapid development of the theory of the electromagnetic field did not introduce fundamental changes in the perfect classical picture.

But science is in constant motion giving rise to forces that overthrow the most stable and firmly rooted theories. The true classical Max Planck made the first breach in the bastion of classical physics.* A breach that was destined to widen and accommodate a whole stream of fresh ideas that Planck himself could never have imagined. To the very end of his long life Planck was apparently not able to accept these ideas fully.

Let us now look into Planck's discovery. You remember that quite unexpectedly to physicists, the rigorous theory of thermal radiation led to obviously absurd results, something like the human body shining brilliantly. In a search to eliminate this flagrant discrepancy between experiment and theory, Planck demonstrated that all difficulties vanish if we assume that atoms emit electromagnetic energy in separate portions, which were called *quanta*. Note that these portions did not in any way follow from the classical electrodynamics of Maxwell. What is more they turned out to be a very unacceptable foreign body in Maxwell's scheme.

Planck's enormous contribution lies in the fact that he was first to realize the necessity of making a logical leap in accepting an assumption that contradicts Maxwell's electrodynamics in order to account for experimental facts. At some point it was necessary to go counter to classical theory. Perhaps there was something in the interaction of light and charges or in the very laws governing electromagnetic waves that was lacking in precision. Planck did not know. He established a fact that he was not able to explain. Meanwhile events developed at a rapid pace.

From the fact that light is emitted in portions, it does not follow that a light ray is discrete in structure. The way Einstein put it was that if beer is always sold in pint bottles that does not mean that beer consists of

---

* Planck's works on the theory of radiation appeared in 1900. The theory of relativity was created in 1905.

indivisible pint portions. Yet experiments in which light knocked electrons out of materials insistently made the point that light is absorbed only in the form of discrete portions. An emitted portion, or quantum, of light energy retains its individuality in subsequent events as well.

This concept was first stated by Einstein in 1905. From the heuristic point of view that he developed, light is always made up of discrete portions possessing energy and momentum. A "bit" of light unexpectedly turned out to be very much like what is ordinarily associated with a particle. These properties of light were called corpuscular (from the word "corpuscles"), and the associated light particle was given the name "photon".

Light with particle properties? All electromagnetic waves with the properties of corpuscles? Can that be possible? Especially since electromagnetic waves were firmly associated with the idea of distributed or smeared matter in space!

Why is it possible for many different people in different places to listen to the broadcast of a single radio station? Because the waves sent out by the station spread out over larger and larger areas. But this generally correct answer touches only on one aspect of the problem.

Now how are we to reconcile this with discrete, or quantum, views? According to quantum concepts, waves are both emitted and absorbed discretely, as quanta. And each one of these portions is indivisible; the receiver either absorbs a whole one or none at all. But we hear the whole broadcast and not little pieces of it.

Of course there is no paradox at all. The energy of a quantum depends on the frequency; it equals the product of the frequency and the famous universal Planck constant, h*. Even for short radiowaves this product is exceedingly small. And so a sufficiently large energy is obtained by having the transmitter send out stupendous quantities of quanta. "Enough for all." The wind blowing in one's face is an analogy. Here huge quantities of molecules give the impression of a soft pressure of air.

---

* We have already encountered this quantity, true, in a slightly different form—$\hbar$. This is for brevity, it means $\hbar = h/2\pi$

However, we do not always get such a smoothing out. Not only instruments in specially devised experiments, but even our own sense organs are capable of detecting individual events. S. Vavilov performed a remarkable series of experiments in which he found that the human eye is an extremely sensitive instrument capable of responding to just a few quanta of light.

We hardly need, here, to enumerate the experiments which demonstrate conclusively that electromagnetic phenomena clearly exhibit both wave properties (those that indicate an obvious continuity) and corpuscular properties (those that just as cogently prove that we are dealing with something discrete).

Here is a very tempting thing. Recall the wind we just mentioned. There (or, more vividly, in sound waves) everything ultimately boils down to the motion of particles—molecules. It is only the overall averaged picture of their movements that creates what is perceived as a wave or the wind. Perhaps the particles of light (photons) fly about like ordinary particles, denser in some spots, and less so in others, forming what we call an electromagnetic wave. That would clarify matters a bit, wouldn't it? It might, but it is not true.

Absolutely unambiguous experiments have demonstrated that wave properties are manifested by even a single photon. One all by itself! This is certainly some food for thought.

But it is only one of a host of riddles that nature offers the scientist.

Wave-particle duality

If the electromagnetic field was always associated with the picture of matter continuously distributed in space (at least before the quantum theory came on the scene), electrons on the contrary were for a long time pictured as minute blobs of matter. This was stressed by the term "particle" itself, which was constantly coupled with "electron". Ultimately, a particle is simply a Newtonian material point. And the electron was no exception to most physicists. In many instances, this view was

extremely helpful. We touched on a number of them when discussing electromagnetic forces in action.

Gradually it was forgotten that many features in the "classical portrait" of the electron had been painted in beforehand. They had become customary and so natural that to reject them was exceedingly painful. Yet it was becoming more and more obvious that this had to be done. Facts were piling up to show that the "classical electron theory", though yielding good qualitative descriptions, was not at all flawless in quantitative matters. In fact the theory even led to strange, paradoxical conclusions. Suffice it to recall the problem of radiation of electromagnetic waves by hot bodies or the fundamental problem of atomic structure.

It was becoming ever clearer that a radical revision of old concepts was due.

In 1923 a young French physicist, Louis de Broglie came out with an idea so paradoxical that many regarded it with a smirk. De Broglie advanced the hypothesis that an electron and any other particle should have wave properties in addition to their corpuscular properties. In other words, the situation that was established for electromagnetic waves was now extended to encompass all types of matter without exception.

The skeptics did not reign for long. In just a short time the highest authority of all—experiment—voted for an electron with wave properties.

It was demonstrated that when electrons are reflected from a crystal they behave in exactly the same manner as respectable waves.

There could no longer be any doubt that both corpuscular and wave properties are associated with matter in all its manifestations. This new concept became known as wave-particle duality. In other words, both light and electrons manifest what would seem to be mutually exclusive properties of particles (corpuscles) and waves.

Now let's take the electron. How can it be a particle and a wave at the same time? Surely that is impossible. We emphatically pointed out the incompatibility of these two images!

Apparantly, we shall have to say that it is impossible, which means—

Which means that when we say that the electron is a wave and a particle, we are only recognizing the fact that the electron, strictly speaking, is neither the one nor the other. In the ordinary sense of the word, it is neither a particle nor a wave (the same goes for the photon as well).

Therefore, when we use these terms, bear in mind that they are to be understood in the sense that the electron can only approximately be described as a particle. Now what does "approximately" mean?

The uncertainty relation

When one speaks of a particle, or material point, we picture a piece of matter located at a specific place at a definite instant of time and moving with a definite velocity. In the language of physicists, this means that we can specify the coordinates and velocity (or momentum—the product of mass and velocity) of the particle with absolute precision.

When we said that the electron can only approximately be regarded as a material point, we had in view that the coordinates and momentum could only be specified approximately, with a certain error. Quantitatively this is expressed by the famous Heisenberg "uncertainty relation".

The Heisenberg relation states the important fact that the more exactly the momentum is defined, the greater is the inaccuracy in determining the coordinates. It would be convenient to write this down in the form of a simple relationship. Let us designate by Δx the indeterminacy of a coordinate, and by Δp the indeterminacy with which the momentum is defined (more precisely, it is the component of momentum along the x-axis). Then the uncertainty relation is written as follows:

$$\Delta p \geqslant \frac{h}{\Delta x},$$

where h is Planck's constant.

A similar relation connects the imprecision of the energy and the indeterminacy of the time interval during

which the process occurs:

$$\Delta E \geqslant \frac{h}{\Delta t} .$$

We give the uncertainty relation without any detailed explanations, for these would require too great an excursion into the theory of microphenomena, and that we cannot undertake. But a few corollaries from the relation we shall examine in somewhat more detail.

Certain corollaries

First of all, let us dispel any illusions that might have arisen. If any "particle" (all objects) have wave properties, why don't we find that a table or the book we are reading or any other of many things we handle possess wave properties?

The answer is simple: because they are heavy. Their mass is great, which means that for an absolutely negligible uncertainty in the velocity, the indeterminacy of the coordinates may be regarded as practically zero. We may consider (not approximately but exactly) that a chunk of something does not exhibit any wave properties at all.

It is only when we deal with small masses, when the objects of investigation are individual elementary particles (or small aggregates of such particles) that uncertainty comes in as a fundamental factor which cannot be ignored.

We cannot disregard the fact that such a concept as path (or trajectory) loses all meaning. We cannot specify both position and velocity at the same time. Therefore, we cannot ignore the fact that the notion of "acceleration" is meaningless.

In a word, we cannot ignore the following fundamental fact: *the Newtonian description of motion becomes impossible*. This is all the more important to us, since, as was stressed, a definition of the concept "force" is rigorous only in Newtonian mechanics. If we are now convinced that the Newtonian description of motion is not applicable to the microworld, then we are forced to draw the following logical and inevitable conclusion: when studying events in the microworld, *we must give up*

forces as a measure of interaction. We mentioned this in passing in the introduction.

Then what have we got left? We have the energy of interaction. How profound and universal is this law of conservation of energy! Energy has turned out to be a much more viable thing than force, and energy takes upon itself the entire load in descriptions of mutual influences in microphenomena.

Interaction and volleyball

Not only the criterion of interaction has changed, the very mechanism itself appears in a new light. You remember how long and persistently the search went on for an intermediary in the interactions of bodies. The search finally ended in the establishment of the concept of a field, the electromagnetic field for one. However, as we have already mentioned, wave-particle duality compels us to seek features of the discrete in the continuous. Now the field has a particle "visage" as well. Hence, it is also possible to interpret interaction from the particle standpoint. Earlier we said that one charge creates a field which acts on a second charge. Now we have the right to say that the first charge creates (emits) quanta (intermediary particles), which are then absorbed by the second charge. This exchange of intermediary particles is the mechanism of interaction and is the way we translate the older classical picture into the quantum language. Whereas before the reaction of one body on another was associated in our minds with some kind of lines or filaments extending from one to the other, now it is appropriate to picture something like a game of volleyball between particles.

However, this new description of interaction is not simply pouring old wine into new wineskins. The quantum interpretation has brought to light numerous fresh opportunities. We shall soon see that this is a real revolution as regards interaction. But before going into these opportunities, let us revert for a moment to the beginning of the book. You recall the argument about close-range action and action at a distance. Just recently—at the start of the century—the very necessity of seeking out a "mediating agent" for interaction seemed

dubious to many. Then the field concept was introduced as a carrier of interaction. But even the field was regarded by many as a kind of substitute intermediary, so great was the gap that separated it from the "real" matter described by Newton's laws of mechanics. Finally, we took another and exceedingly important step, we found out for sure that the intermediary was not only material (having energy, momentum, and so forth) but could justly (and precisely) be regarded as particles, exactly like the sources between which the interaction took place. There was no gulf after all. Both the interactors and the carriers of the interaction turned out to be ordinary matter, simple elementary "particles", the quotes calling to mind the great distance that science had traversed from Newtonian motion to wave-particle duality.

Now what are these new fresh opportunities?

Here is a question: Do the particles of the electromagnetic field have a monopoly on transferring interaction? Perhaps other particles (or groups of particles) could take upon themselves the role of carriers of interaction.

This is a rather interesting and fruitful idea. We shall come back to it in the next chapter. For the time being only two important factors will be examined.

The charge has a face-lifting

The first has to do with the charge, our old friend, the electric charge. The larger it is, the greater its effect on surrounding charged particles. In quantum parlance, this means that the larger the charge, the more quanta (interaction carriers) are sent out in all directions by the source. Thus, we can now say that a charge is a criterion of activity, of intensity of emission and absorption, a source of mediating quanta.

If these quanta are the quanta of the electromagnetic field, then the charge is an electric charge. But, as already mentioned, other particles too can be mediating quanta. Hence the necessity to introduce other types of charge as well. Each type of mediating particle has its own charge and its own coupling constant.

This is a very important conclusion!

We can go through the table of elementary particles and try each one out to see if the particular particle (or group of particles) fits as a mediating agent. The only constraint is that it should not come in conflict with the conservation laws. But nature imposes additional limitations, so that we really do not get the extreme diversity of interactions that one would expect. There are very few different types of charges. That is actually our task: to pick out all known types. We shall come back to this again in the chapter on weak interactions.

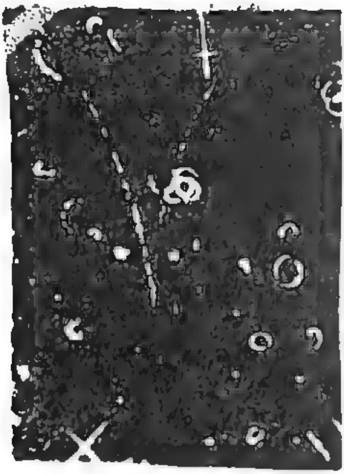
Reincarnation in the world of the ultrasmall

Our discussion about different charges led us on to a topic that is extremely important all by itself. A few lines back we wrote that a particle (we have in mind one that carries interaction) is emitted by a source. But what does "emit" mean here? At any rate it is not like opening a cage door and letting out a canary. The point is that there was no such particle inside the source prior to emission, it was not being kept in reserve. A photon is not hidden in the atom but is born in the very act of emission.

"Born" is the word!

Hence, particles can be created (and, we can add, annihilated too). That is precisely the conclusion that we arrive at by the foregoing chain of reasoning. But perhaps we have hurried with the conclusions. It may be that the photon is some kind of anomaly, a nontypical particle (and this is substantiated by the fact that what we know as ordinary matter cannot be made out of photons).

These doubts were justified up until 1927, the memorable year in which a young English theoretical physicist, Dirac, published a paper where he began by trying to write the equation of motion of the electron that would satisfy the requirements of relativity theory. At first glance, a rather routine problem. However, it soon became apparent (true, not without a certain amount of "friction" and experimental stimulus) that this is pos-



sible only when we assume the electron to have a partner particle in every way identical to the electron but differing in the sign of the charge. It was actually detected in a cloud chamber and was christened positron. Like the electron, this particle, taken separately, is entirely stable—it can exist for any length of time. However, the theory predicts that when an electron and positron meet, they annihilate each other giving rise to high-energy photons (gamma-quanta). The reverse process is also possible—the generation of an electron-positron pair (separate electrons are not generated because this would violate the law of conservation of charge). This occurs, for example, in the collision of a gamma-quantum of sufficient energy with a nucleus. In a cloud chamber placed in a magnetic field, such a pair leaves a characteristic track in the form of a two-prong fork.

The electron, the oldest of the particles and the most important building material for numberless atoms the reliable and experienced electron was not eternal after all. It could vanish! And it could reappear again! This fact was startling to physicists and—after experiment brilliantly corroborated the predictions of theory—it gave rise, said one famous theoretician, to a fantastic complacency. It was a long, long time since theory had seemed so powerful and all the secrets of nature so accessible.

The works of Dirac do indeed occupy a very significant place in modern physics. No wonder then that their author was surrounded by an aura of fame.

When the authors of this book were students, physicists had a favourite song they called "The Song of the Electron".

Oh, Apollo, handsome patron of the arts,
Inspire the heart of the singer,
And in the name of the great son of Lato and Zeus let me sing of love—
Mother of beauty and sister of immortality!

Of youthful Eros eternally and of foam-born Aphrodite.
Let my song ring out in words the truth
Of how, perturbing the ether,
With velocity less than the speed of light,
And uniformly without any exterior field
The *Electron*—so called by the glorious gods—flew.

There were many electrons like this one in flight,
But obeying Coulomb, they made small approach and then
Rapidly went apart. And the reason was that only
At rest could they repel one another,

But never attract!

And thus this electron that flew perturbing the ether
With a velocity less than the speed of light,
Uniformly without any exterior field, complained!
It complained to god Dirac and ardently prayed
in the following words:

"Oh God almighty, oh God invariantly* great,
Who inserted a third point in my equation**
Why did you despise the laws of the universe
And order me to wander about and never know
the joy of love?
Give me a mate!"

Such was the passionate prayer the electron made to
the terrible god Dirac.

The god furrowed his brows and triumphantly stated
"Let it be!"

---

* Invariant means the same in all systems of reference.

** The earlier mentioned force of radiation friction due to interaction of a point charge with its own electromagnetic field. This force is expressed in terms of the third derivative of the coordinates with respect to time.

And he hurled into a nucleus a photon f $h\nu$*
Greater than $2mc^2$,
Heralding a new era with the birth of the first pair!

And from then on, every electron, even
The outermost, just like you and me, honoured reader,
Has no cause for sorrow, for it has a true mate!

And like yours, she is fair with a level of energy less than zero,
We cannot say that her character is upright and stable,
But in positivity, it is far above anything
The electron will ever achieve in all time!

Ending this song, I want it to ring out as a
Hymn to Eros, the handsome ubiquitous god of love
That dwells in the realm of the gods and in the hut of the shepherd.
That lives in the breath of the air, in the blooming flowers,
Asserting the eternal and beautiful law of the universe:
Each ψ will meet its ψ * !**

One of the copies of the song bears the autograph of Paul Adrien Maurice Dirac.

The idea of particles and antiparticles proved fruitful in the extreme. All particles were found to have partners (true, in exceptional cases, like the photon, the particle and antiparticle simply coincide). The antiproton, antineutron, and so forth have been found experimentally. We now know that generation of pairs and annihilation are not the monopoly of electrons and positrons.

What is more, it was realized that mutual transformations, that is, the annihilation of certain particles and the generation of others, does not at all necessarily have to proceed via generation of pairs—particles and antiparticles and their annihilation. Elementary particles have extremely diverse reactions (this term was borrowed from chemical parlance and is very fitting). Yet there is much in common too. The collision of particles is like the blow of steel on flint. The flint is the bombarding particle with high energy. Again particles or groups of

---

* Read "h nu"—the product of Planck's constant by the frequency, which is the photon energy.

** Read "psi" and "psi conjugate" which usually designate the "wave function", the quantity used in quantum mechanics to describe particles.

particles serve as the target or "flint". Each blow brings out "sparks"—new particles. And it breaks down both the flint and the steel. The stronger the blow, the more sparks, or particles. At times there are several hundreds of particles. Today we have extensive experimental material on the generation of pairs of particles. This mass of information leaves not the shadow of a doubt: all particles without exception can vanish and reappear.

But doesn't this run counter to the most fundamental of all the laws of nature—the law of conservation of matter? How can matter vanish, turn into nothing, and appear out of nowhere?

Naturally, we didn't try to say that at all. In spring, when the bare branches of the trees produce buds and then leaves, and when in autumn they bend under the heavy weight of fruit, does anyone ever imagine that this represents a break with the law of conservation of matter? Buds, leaves, and fruit do not appear out of nothing. It is simply yet another instance of the numberless links in the eternal cycle and mutual transformations of matter in nature.

When dealing with the generation and annihilation of particles, we encounter a very unusual yet ultimately profoundly related transition of matter from one form into another. One may speak here of the transition of matter from one state to another.

In the annihilation of an electron and a positron, matter passes from the electron-positron form to the electromagnetic form. There is of course no vanishing at all. Incidentally, in this and all other processes the charge, energy, momentum, etc., are conserved, thus once again compelling us to regard all these phenomena simply as transformations.

The obvious and the hidden

What the foregoing suggests is that matter should be described in a unified manner, and the various particles are to be regarded as diverse manifestations of this common matter. An enticing idea! And though attempts to construct such a universal theory have been made,

we find it hard to see any radical advance. We are as yet unable to "build" particles. So we have to confine ourselves to a superficial description. Our situation is somewhat like that of a botanist studying the life of a plant from a number of photographs depicting the seed, then the shoot, then the flower, and, finally, the seed again. The botanist would then be sure of one thing: there are different states of the plant: seed, shoot, flower. He would also know that they follow one another in a sequence, which would allow him to speak of certain laws of transformation. But these photographs would hardly allow him to establish any "inner dynamics" of the phenomena.

The physicist also has a series of photographs depicting what we rather arbitrarily call "elementary particles". The term is justified today in the sense that we hardly know anything about the structure of these particles.

All elementary particles are introduced into the theory directly from experiment. But don't regard this in too narrow a sense, because we include here the values of charge, mass, spin and so forth, and also subtle features of the laws of motion. This is not accidental, for theory develops on the basis of experimentation. Now the experiment, in general outline, looks something like this: the recording device (a cloud chamber, photographic plate, system of counters, etc.) exhibit one or several tracks of beams of primary particles. All the fine details of interaction are hidden from the observer, who sees only the result of the interaction—again in the form of tracks of secondary particles. Not everything is so simple, however. Certain of the participants have no charge and therefore do not even leave traces of their whereabouts. But what we wanted to say was that experiment only yields indirect evidence, on the basis of which we conjure up a picture of interactions.

A comparison. Not too bold?

Since we do not see the actual interactions but only their results (the transformation of particles into other particles or into the same particles but in a different

state), it is quite natural for us to strive to reflect this situation in theory. Then as a corollary we get a physical picture with particles as the central feature and as something obtained directly from experiment. Let us try an analogy. Imagine for a moment that we know nothing about the molecular structure of matter. Then even such a simple problem as the melting of ice would probably appear in the following light. Researchers would provide us with a detailed list of the properties of ice and of water. Studied experimentally, don't forget. They might even go so far as to call ice an "elementary" entity, and water another "elementary" entity. Then, relying completely on experiment as before, they would formulate the following law: under specific conditions (temperature and pressure in this case) ice passes into water.

But how? What internal hidden changes take place? Without the molecular picture this cannot be determined. And here our imaginary scientists would be in the position we mentioned where they cannot grasp the "inner dynamics" of the process. In the fundamental problem of elementary particles, the main thing is that we do not know the structure (the inner laws) of these particles. That precisely is what compels us to regard them as integral entities and describe the multiplicity of processes in the microworld only as a vanishing of particles and a generation of new ready-made particles.

This method is not so bad however. Physicists are experienced pathfinders. They have deciphered particle tracks, unearthing hidden effects that give away peculiarities of particle behaviour. We now know not only the laws of motion of free particles, but a great deal as well about their interactions. As we have already mentioned, on modern views, it amounts to a particle exchanging with another particle quanta of the mediating field. In other words, the exchange is that of particles, but particles of a different nature.

The nature of the quanta emitted and absorbed by a particle is determined by the type of charge of the particle. If it is electrically charged, then it has the right to emit and absorb photons. If it has a nuclear

charge (we'll talk about this charge later on), then pi-mesons, and so forth.*

Each such fact of emission or absorption transfers a particle from one state to another.

Interaction with vacuum

We have said that interaction consists in the fact that one particle emits quanta and another particle absorbs them. Now, can a particle absorb the quanta that it emits? Well, there is no reason why it can't. Such processes result in the interaction of a particle with itself. This is sometimes interpreted differently: as the interaction of the particle with the vacuum. Though it may sound paradoxical, the expression is fully justified. The point is that when we speak of self-action, we mean that the particle is acted upon even when it is all alone, when there is nothing around it, not a single "real" particle. In other words, when a vacuum or void surrounds it.

The vacuum plays an important role in physics. The textbooks speak of "polarization of the vacuum", "vacuum corrections", "vacuum oscillations" and so forth. Yet just a short time ago it was considered absurd to speak of the "properties of a vacuum". What kind of properties can emptiness have? Properties are something inherent in matter, but vacuum means no matter...

Stop! That's enough. This is the crux of the matter. You say "no matter". That means "not a single particle". This is not so simple after all. In spring when the grass begins to grow, how do we answer the question: is there any grass? "It will be there when it becomes green, when shoots come out of the ground," you may say. And before that? When the shoots have not come up yet, when they are still in a "subterranean existence"?

This is getting to be rather scholastic, isn't it? "Already grass" or "not yet grass". True as that may be,

---

* Nothing here conflicts with the law of conservation of energy: the duration of the processes, Δt, is very small and, according to the uncertainty relation, energy "smearing" should be very great.

the point is this, however. We can see grass that has grown, we can perceive it with our sense organs. But there is a bare field where the shoots have not come through yet—a kind of "vacuum" or emptiness in a certain sense of the word.

This is no far-fetched analogy. In the theory of elementary particles, a vacuum may be regarded not as a void, a nothingness, but as a peculiar state of all particles that have so little energy that they cannot be perceived in any way either by the human sense organs or by the most delicate instruments. But vacuum particles "feel" the influence of "real" particles. They rearrange themselves due to the presence of real particles, thus, incidentally, resulting in experimentally observable effects. *If the effect is sufficiently energetic, then the particle is transferred from the "invisible vacuum state" to the ordinary real state. Outwardly, this appears as the generation of a particle* (*its birth*). *In exactly the same way, we can picture the annihilation of particles as their transition into the vacuum state.*

This mode of description is not only possible, but is quite natural in present-day theory, for it permits of a simple description of the processes of generation and annihilation of particles (without going into their inner dynamics, as we have already said) by reducing all processes to transitions from one state to another (true, one of the states is rather exotic). By introducing the vacuum concept and even building a "vacuum theory", physicists have made considerable advances not only in "ordering their theoretical affairs" but also—and most importantly—in describing experimental facts. We now compute more precisely the energy levels in atoms, substantial corrections have been found for the values of the magnetic moments of electrons, and so forth.

But the "vacuum problem" has again brought forth the great difficulties of the theory. A whole series of quantities, such as the energy of interaction of a particle with itself (or with the vacuum, which is the same thing), vacuum corrections for charges, and so forth, become infinitely great, which of course is a depressingly absurd result. One thing is clear, patches won't help the matter. A fresh grasp of the nature of elementary par-

ticles will probably force us to reconsider much of what now (already!) seems customary and natural.

Meanwhile... a vacuum is pictured as an endless sea permeating everything, out of which particles jump up here and there like monsters.

The Sea of the Unknown, one might say.

Of infinite depth.

Yes, a whole new book is needed to look deeper into this unknown.

CHAPTER FOUR

Mighty forces
Fused into worlds.
And a marvellous, beautiful
Life emerged.

The Great Word (A. Koltsov)

NUCLEAR FORCES

1

On the boundary of the unknown

The atomic nucleus, nuclear energy, the atom age. These and dozens of other terms with the word "nucleus" appear in newspaper articles, books and scientific papers, and give rise to fear and hope. No scientific discoveries ever meant so much to humanity as those of nuclear physics. Even people far removed from physics come into the discussion.

At the same time the researcher sees before him blank spots galore. But how is this possible when atomic power stations are generating electricity, a nuclear icebreaker is plowing its way through the Arctic waters, and nuclear specialists are at work in every imaginable field, from metallurgy to the production of Christmas-tree decorations. Is it possible that a science which has achieved so much can still have obscurities (and no small number

of them) in its very foundations? There isn't really anything paradoxical. We are rather like a bricklayer who knows how to lay bricks and build houses but does not know all the properties of the bricks he is laying, and practically knows nothing about the way the bricks are manufactured. In such cases we often say that we have studied some of the properties but know nothing about the essence. This may not be the most suitable expression, but it shows that we are not able to explain a great mass of experimental data in a unified manner. And don't be too surprised, because nuclear physics possess problems whose solution involves the ultimate problem of the structure of matter, the structure of the elementary particles. To continue the analogy, these particles are often called the building blocks of the universe. Here we find ourselves much in the position of Carolus Linnaeus: we simply haven't advanced much beyond the limits of systematics. Practically nothing is known about the structure of these particles and even about what should be understood by the term "elementary".

We are on the borderline of ignorance, the unknown. The boundary is unstable and has been attacked time and again, but there has been no radical breakthrough.

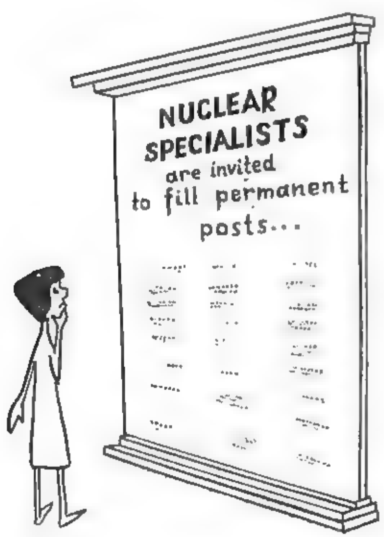


TABLE OF ELEMENTARY PARTICLES

	Name	Symbol		Mass	Spin	Electric charge	Lifetime, sec	Decay products
		particle	antiparticle					
	Photon	γ	γ	0	1	0	Stable	
Leptons	Electron neutrino	ν_e	$\bar{\nu}_e$	0	$\frac{1}{2}$	0	Stable	
	Mu-meson neutrino	ν_μ	$\bar{\nu}_\mu$	0	$\frac{1}{2}$	0	Stable	
	Electron	e^-	e^+	1	$\frac{1}{2}$	-1	Stable	
	Mu-meson	μ^-	μ^+	206.7	$\frac{1}{2}$	-1	$2.2 \cdot 10^{-6}$	$e^+ + \nu_\mu + \bar{\nu}_e$
Mesons	Pi-mesons	π^0	π^0	264.2	0	0	$2.2 \cdot 10^{-16}$	$2\gamma,\ \gamma + e^+ + e^-$
		π^+	π^-	273.2	0	1	$2.6 \cdot 10^{-8}$	$\mu^+ + \nu_\mu$
	K-mesons	K^+	K^-	966.5	0	1	$1.2 \cdot 10^{-8}$	$e^+ + \nu_e + \pi^0$ $\mu^+ + \nu_\mu,\ \pi^+ + \pi^0$ $3\pi,\ \mu^+ + \nu_\mu + \pi^0$
		K^0	$\bar{K}^0$	974	0	0	$K_1^0\ 1.0 \cdot 10^{-8}$ $K_2^0\ 6 \cdot 10^{-8}$	$\pi^+ + \pi^-,\ 2\pi^0$ $\pi^0 + \pi^+ + \pi^-,\ \pi^+ + e^- + \bar{\nu}_e$

Continued

	Name	Symbol: particle	Symbol: antiparticle	Mass	Spin	Electric charge	Lifetime, sec	Decay products
Baryons	Proton	p	$\bar{p}$	1836.1	$\frac{1}{2}$	1	Stable	
	Neutron	n	$\bar{n}$	1838.5	$\frac{1}{2}$	0	1013	$p+e^-+\bar{\nu}_e$
	Lambda hyperon	Λ^0	$\bar{\Lambda}^0$	2182	$\frac{1}{2}$	0	$2.5\cdot10^{-10}$	$p+\pi^-$, $n+\pi^0$
	Sigma hyperons	Σ^+	$\bar{\Sigma}^+$	2327	$\frac{1}{2}$	1	$0.8\cdot10^{-10}$	$p+\pi^0$, $n+\pi^+$
		Σ^0	$\bar{\Sigma}^0$	2331	$\frac{1}{2}$	0	10^{-11}	$\Lambda+\gamma$
		Σ^-	$\bar{\Sigma}^-$	2340	$\frac{1}{2}$	-1	$1.6\cdot10^{-10}$	$n+\pi^-$
	Xi-hyperons	Ξ^0	$\bar{\Xi}^0$	2565	$\frac{1}{2}$	0	$1.5\cdot10^{-10}$	$\Lambda+\pi^0$
		Ξ^-	$\bar{\Xi}^-$	2580	$\frac{1}{2}$	-1	$1.2\cdot10^{-10}$	$\Lambda+\pi^-$

NOTE. The mass, spin and lifetimes of antiparticles are the same as for the particles. The charge of an antiparticle is opposite in sign and equal in magnitude to that of the particle.

For that matter, nuclear physics does not differ from any other science in this respect. Just a little more thought, and the series of why's will bring us to a realm that has not even been studied. How right the saying is that when studying some problem, man passes through three stages: one, "everything's clear", two, "nothing's clear", and three. "certain things are clearing up".

What are nuclei made of?

After all, why does the atomic nucleus compel us to study the elementary particles? Take planetary motion. The planets are ultimately made up of elementary particles, yet we need not emphasize them in any way.

The obvious reason is that in atomic nuclei there are so few particles that the properties of each one separately do not average out, but play a decisive role.

Which means that though we want to build a structure, we shall have to deal with the bricks first. This is all the more important since we cannot discuss intranuclear forces without first looking into the composition of nuclei. Here again, nuclear physics confronts us with a new situation.

Indeed, neither gravitational nor (though to a lesser degree) electromagnetic forces required a detailed story about the structure composition of the pieces of matter that participate in the interaction. Nuclear matter is so peculiar that it is impossible to separate the question of what is interacting from the question of how the interaction occurs.

One little girl said that a hammock is "a lot of knots connected with pieces of rope".

We could say the same thing about the atom—a nucleus and lots of electrons and electric fields to hold the particles in place. And when speaking about the constitution of the atom we disregard the fields in accord with a tradition to bring out *what* is bound and forget about the *thing* that does the binding (the roots of this approach go back into mechanics). But in the nucleus the situation is radically different. Here the knots are

somehow inextricably linked with the ropes. That is why we recalled the girl and the hammock.

It is now probably clear why we start with the composition of nuclei before talking about the nuclear forces.

Physicists are acquainted with about thirty* types of more or less stable elementary particles. They differ in mass, electric charge, and other intrinsic properties. Quite a collection to choose from for the building material of atomic nuclei. Suppose we have two tables, one of nuclei and the other of the elementary particles. Speaking of mass, the hydrogen atom has the lightest nucleus**. It is 1836.12 times heavier than the electron and has an equal and opposite (positive) charge. Of the elementary particles, only one (the proton) has exactly the same properties. So we have the composition of one nucleus. But it will not be so easy with the other nuclei. Take the element next to hydrogen in the Periodic Table—helium. The helium nucleus (we shall disregard isotopes for the time being) is almost exactly four times heavier than the hydrogen nucleus. Perhaps it consists of four protons? But then its electric charge would be four times the proton charge. In actuality it is only twice as large. It might be possible to eliminate this difficulty by presuming the existence in the nucleus of other particles than the proton. They would be charged negatively and would compensate for the extra charge. And if these particles have small masses then we might make ends meet. At first glance this appears attractive, all the more so since we have our old friend the electron as a suitable particle. It does appear attractive, but theoreticians and experimenters were instantly against this electron-proton metal. Their main reason was that the electron is too lightweight. We'll go into that later on.

For the present, it is clear from the table that there is not such a big choice after all. Take a look at the lifetime column. It ranges from roughly a thousand seconds in the case of the neutron to the fantastically small interval of time of $2.3 \pm 0.8 \times 10^{-16}$ sec for the

---

* The number is not exact because even the criterion of "elementariness" is still very obscure.

** The mass of the atom practically coincides with that of the nucleus. The electron accounts at best for only about five hundredths of one per cent.

pi-zero mason (π^0). When their lifetimes are up, the particles decay into other particles.

But atoms, and their nuclei too naturally (say the helium nucleus) do not decay by themselves; in fact to break them up is no simple job at all. They are stable and, hence, should consist of stable particles only. But there is not a single stable elementary particle except the proton and antiproton (not counting the light particles, which, as we know, cannot "get along" in the nucleus). Not a single one!

Now what have we got? Nuclei cannot consist of protons alone, that is clear. The remaining particles are either too light to be component nuclear parts or are unstable. Where is the way out?

On common sense

One thing is definite: we would be totally helpless confronted with the mysteries of the nucleus if it were not for quantum mechanics. This is the kingdom of microphysics in the fullest sense of the word. Things are all too often paradoxical here if judged from the viewpoint of our customary notions developed in a world of "large things". Our intuition based on classical entities is frequently more an enemy of investigators than an ally.

Let us start with the circumstance that the light particles (at any rate, electrons) cannot be nuclear constituents. The classical theory explains nothing. But with the uncertainty relation, things fall into place.

Nuclei have very small dimensions. Numerous experiments have shown that these dimensions are of the order of about one hundred thousand millionths of a millimetre. The latter, therefore, is the uncertainty of the coordinates of an intranuclear particle. This straightway enables us to determine the uncertainty of momentum and, consequently, the velocity (because the mass of the particle is known).

We take another step: we recall that the kinematic energy is equal to one half the product of the mass by the square of the velocity and so we can find the spread in energy values. It will immediately be seen that it is inversely proportional to the mass of the particle. For

heavy particles such as protons the spread is comparatively small, but for electrons it increases about two thousand times over and becomes appreciably greater than the experimentally found nuclear binding energy, the energy with which nuclear particles interact. But if the binding energy is less than the kinetic energy, the forces of interaction are insufficient to hold a particle, which will very soon overcome any restraining force and leave the nucleus.

Thus, even if a light particle somehow gets into a nucleus, energy considerations alone are sufficient proof that it will not remain there for long.

About composition again

And so we must look for the building materials of nuclei only among the heavy particles. In addition to the proton, there are quite a number of suitable particles in the latest table: first of all the neutron and a large group of so-called hyperons.*

Generally speaking, hyperons can enter into the composition of a nucleus. They go to build what are called hyper-nuclei, which have comparatively recently been found experimentally. Hyper-nuclei however are not stable, disintegrating very rapidly, which is not surprising since the hyperons themselves have lifetimes that do not exceed one ten-millionth of a second.

Which leaves us with only one particle, the neutron. Physicists have known the neutron for quite some time. It was discovered by a young scientist, Chadwick, in Rutherford's laboratory in 1932.

The neutron has no electric charge. Its mass is just about that of the proton (the proton is 1836 times heavier than the electron and the neutron is 1839 times heavier, which is hardly any difference at all).

---

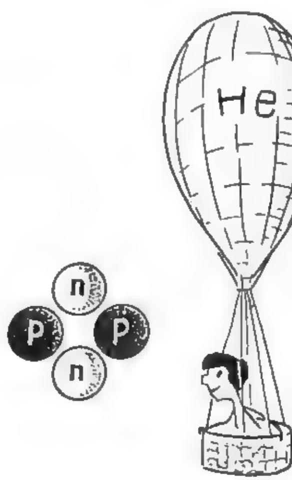
* We shall not discuss antiprotons, antineutrons, and other antiparticles. When an antiparticle encounters a particle (say, an antiproton and a proton)—such encounters are inevitable in our world—the result is what physicists term annihilation. The particle pair vanishes and turns into new particles. Due to annihilation, a nucleus consisting of both particles and antiparticles cannot exist for long.

The proton-neutron model

A heavy neutral particle—perhaps that is what goes to make up the nucleus along with the proton. Let's take the helium nucleus. It has a charge twice the proton charge, and a mass almost exactly four times greater. If we assume this nucleus to consist of two protons and two neutrons, we will have just what we are looking for. Excellent results are also obtained for the nuclei of all other elements. Not only the charge and the mass but all the other characteristics are likewise in perfect agreement with experiment.

The neutron seemed so fit for the role of nuclear particle that at least in two countries (in the Soviet Union—Ivanenko, Gapon—and in Germany—Heisenberg) theories were advanced at practically the same time, right after Chadwick's experiments. They formulated the basic ideas of the proton-neutron model of the nucleus, the model still generally recognized.

However, there was still the problem of reconciling the stability of nuclei on the one hand with the instability of the neutron on the other. Though rather stable on the background of other particles, the neutron does disintegrate after all (in about 12 minutes). How is one



to account for the established fact that dozens of types of nuclei live longer than 12 minutes, while a very large proportion have practically eternal lifetimes?

The immutable in the mutable

Now just what is stability? The still water surface of a pond in the woods? Here apparent immobility and stability hides intense motion. Molecules of water are constantly evaporating from the surface and, conversely, molecules of vapour are constantly being captured by the water. If these two counter streams of molecules are the same, the water level does not change—we have equilibrium.

Stability does not mean a total absence of motion. In most cases there is motion. The important thing is that the character of motion constantly restores the system. In such cases, physicists speak of dynamic equilibrium.

But is the stability of a nucleus precisely this type of dynamic stability? It obviously is, because there does not appear to be any other possibility.

The question is: How is dynamic equilibrium maintained? Apparently, the neutron inside the nucleus participates in some kind of processes, on the background of which its instability ceases to play any role.

What are these processes?

2

A comparison to which we shall return

By way of an illustration frequently used to help picture intranuclear dynamics, imagine that two men are carrying a load, a heavy load that is too much for a single

person but too unwieldy for two to handle. Another restriction is that the load should not be dropped to the ground, and no time is allowed for resting. What is more, if it does drop, it cannot be picked up again.

If it weren't for the second man, the first would sooner or later drop the load. Now let's come back to the neutron which when alone disintegrates. The load can be carried by two men changing over every once in a while.

Perhaps something like this happens in the nucleus. It would seem that the presence of the proton alongside the neutron stabilizes the latter, for otherwise the free, separate neutron inevitably disintegrates, yet in the nucleus it behaves as a completely stable particle.

To come back to our illustration again. What can play the part of the load exchanged between the proton and the neutron?

Everything reduces to interaction after all

Let us think hard for a moment and figure out what considerations support this analogy. Do we necessarily have to assume that the neutron and proton exchange something? And, finally, what is the nature of this "something"?

We have now come to a point where we can recall an important detail: namely, that the particles in a nucleus do not only co-exist, they are intimately interconnected. It is not enough simply to stabilize the neu-

trons, it is also necessary to account for the stability of the entire nucleus. These two problems, it appears, are linked up very closely, indeed.

We have thus approached the question of intranuclear interactions.

A few pages back we found volleyball helpful in illustrating interactions in quantum theory. Particles bandy about (exchange) quanta of some "mediating field". On this view, the interaction of protons and neutrons within the nucleus should depend on the type of exchange particles. Particles that convey the interaction.

This physical picture is pictorial and convincing, just like the two men and the heavy stone, the point being that the only way they could carry the load was by constantly handing it over to one another. But to do this they had to be close together. The necessity (and possibility) of exchange binds particles. To an outside observer, it will appear as if attractive forces are operating.

This is of course a comparison, but all analogies serve the same purpose: to construct a vivid image to facilitate a more profound understanding of the phenomenon at hand.

So we again come to the question of "mediating particles" that cement the nucleus together. What are these particles and what properties do they have?

Classical mechanics and the nucleus

Again we repeat for the umpteenth time that without quantum theory we would be helpless. Indeed, let us picture for a moment that the nucleus lives according to the laws of classical physics.

Let us take the most elementary of nuclei, the deuteron. It consists of one proton and one neutron. Here they are standing side by side eager to interact, that is to say, exchange some sort of particles. But classical mechanics prohibits this, because to initiate interaction, each particle must eject (emit) and absorb particles.

Within the framework of classical mechanics, the laws of conservation of energy and momentum prohibit a free

particle from participating in any kind of emission (say, a free electron cannot emit or absorb electromagnetic waves). It is important not to confuse the emission of a particle with disintegration. When a neutron emits some kind of particles (we designate them by A) the process occurs as follows:

$$\text{neutron} \rightarrow \text{neutron} + A.$$

In other words, the neutron remains intact after the transformation.

The situation is particularly transparent when dealing with mass. The point is that if a proton (or neutron) emitted some particle, that particle would carry off a certain amount of mass. Now, running ahead a bit and locating in the table of particles the mass of the carriers of intranuclear interaction, we can subtract it from the mass of the proton and see that the remainder does not fit any of the masses in the particle table. Paradoxical indeed! Surely it can't be that a proton or neutron ejects some mediating particle and then turns into something that does not exist!

Yet this is not the only paradox. Strictly speaking, all the phenomena inside a nucleus are absolutely paradoxical if approached with the yardstick of classical theory. We have already encountered paradoxes of this nature.

Now let us attack the mediating-particle problem from the standpoint of a quantum description of events.

Conclusions that stem from uncertainty

The first to fall is the objection that the law of conservation of energy and momentum prohibits nuclear protons and neutrons from emitting and absorbing any particles. This is due to the fact that, as already pointed out, neither a coordinate and momentum, nor energy and lifetime of any of the constituent particles of the nucleus can have definite values at the same time. The spread, or quantum uncertainty of these quantities to put it in the lingo of physicists, immediately removes these difficulties.

But this is not all. There still remains the paradox of masses. And here we are suddenly aware that quantum mechanics not only saves the situation but even utilizes this "difficulty" to extract fresh information about the quanta which convey the interaction.

However, let us go a bit more slowly and do some calculating. We spoke about the energy spread of particles in the nucleus. Let us concentrate on some proton and denote the energy spread by $\Delta\varepsilon$. Obviously the energy of the quantum (the carrier of the interaction which we denote by E) should fit into the range of this spread, thus allowing us to write the equation:

$$\Delta\varepsilon = E.$$

And now we must take into account the well-known fact discovered by Einstein that the mass and energy are connected by a remarkably universal relation, which states that the energy is equal to the product of the mass times the square of the velocity of light. This is the famous formula:

$$E = mc^2.$$

Yet another step must be taken. What is the value of the energy spread $\Delta\varepsilon$? The uncertainty relation can help us here. As we know, the uncertainty of the energy is connected with the time during which the process occurs by the relationship

$$\Delta\varepsilon = \frac{h}{\Delta t}.$$

What kind of time is this Δt? Obviously, it can simply be equated to the time of flight of the particle-carrier of the interaction. This is the interval between the instant of emission and the instant of absorption of the quantum, that is, it is precisely what might be termed the "time of interaction".

But the time of flight is equal to the distance covered, l_0, divided by the velocity.

Here we are interested only in a qualitative estimate. And so we can take it that l_0 coincides with the dimensions of the nucleus (that is, that each quantum traverses the nucleus from one end to the other) and the velocity

is equal to that of light. Then we get

$$\Delta t = \frac{l_0}{c}.$$

Now let us bring together all our equations:

$$\Delta\varepsilon = E,$$

$$E = mc^2,$$

$$\Delta\varepsilon = \frac{h}{\Delta t},$$

$$\Delta t = \frac{l_0}{c}.$$

It is now easy to find the mass of the mediating particle:

$$m = \frac{h}{l_0 c}.$$

It is a remarkable fact that in this equation, all the quantities in terms of which m is expressed have long since been found from experiment. Substituting the values of the Planck constant h, the dimensions of the nucleus (or, more precisely, the "radius of interaction") l_0 and the velocity of light c, we find that m should come out to roughly two or three hundred electron masses.*

The reader will have to forgive us this mass of calculations, but we must say that the magnificent result obtained was really worth the trouble.

We were able to elucidate some very essential features of nuclear interaction:

1. Interaction is the result of an exchange of particles.

---

* The same result may be obtained from the uncertainty relation for a coordinate and the momentum:

$$\Delta x \approx 10^{-13} \text{ cm}; \qquad \Delta p = \frac{h}{\Delta x}.$$

Computing the uncertainty of the energy of the neutron or proton in terms of the uncertainty of the momentum $\Delta\varepsilon = \frac{(\Delta p)^2}{2M}$ (M is the mass of these particles) and taking into account that $\Delta\varepsilon = mc^2$, we get an expression for the mass of the quantum that conveys the interaction:

$$m = \frac{(\Delta p)^2}{2Mc^2} = \left(\frac{h^2}{\Delta x}\right)^2 \Big/ 2Mc^2.$$

Substituting numbers, we get just about the same result as by the other method.

2. The distance over which the interaction occurs (or the range, as it is called) is the smaller, the greater the mass of the particles carrying the interaction: $l_0 \sim \frac{h}{mc}$.

3. The interaction is specifically quantum interaction (Planck's constant h is present).

The meson: first discovered in theory

These exceedingly interesting conclusions were first drawn by the Japanese scientist Yukawa. At that time the list of elementary particles was very modest: the photon (the quantum of the electromagnetic field), the electron with its mirror image the positron, the proton and the neutron. And that, actually, was all. With remarkable scientific audacity, Yukawa, after a thorough analysis of the facts, stated that there must be a new particle, with a mass approximately two hundred times that of the electron for handling intranuclear interactions.

This prediction was brilliantly confirmed. The particle which Yukawa called the meson (actually there were three such particles with only slightly different masses but electrically positive, negative and neutral) was soon discovered experimentally and the properties conformed exactly to those required by theory. The meson theory of nuclear forces accounts for many aspects of this phenomenon.

Short-range action

These forces act over very small distances. We proceeded from this fact when, on the ultimate basis of experimental facts, we sought the mass of the meson. A similar thing occurs when, say, we break a piece of chalk and then attempt to put the two halves together again by pressing them hard. It doesn't work because the molecules at the broken edges are just a tiny bit farther apart than in the solid piece—which is sufficient to eliminate practically all interaction. In the nucleus this is still more pronounced.

Physicists say that nuclear forces are short-range forces. You can come up close to the nucleus and never feel them, though inside the nucleus the interaction is fantastically strong.

How strong are nuclear forces?

There are tremendous forces at work, stupendous energies raging inside the atomic nucleus. Let us try a comparison with something we are familiar with. Say the energy released when a person sneezes. This is small indeed. Or the work done in picking up a coin from the floor. Every day we do far more work than that.

You will probably be surprised to learn that the work in the foregoing examples is thousands of millions of times greater than the energy required to tear a particle out of the strongest nucleus. Thousands of millions of times!

But then why so much talk about tremendous nuclear energies? Why are gigantic atom-smashing machines constructed that consume the electricity of a whole city to run if a simple sneeze performs more work than is needed to disrupt the bonds inside a large, large number of nuclei?

You have probably guessed the secret by now. The important thing is not the total energy but the amount per nucleus or, still better, per particle of the nucleus. Though enormous energy is expended in lifting a coin—an energy thousands of millions of times greater than the nuclear binding energy, the proportion per nuclear particle is an ultra-minute dosage: less than one millionth of one millionth of the binding energy. And even if we accelerated the coin to cosmic velocities of tens of thousands of kilometres per hour, the energy associated with this motion per particle would still be thousands of millions of times less than the intranuclear energy. That is why it is so hard to obtain a nucleus-smashing projectile. It has to possess energy sufficient to break up a nucleus.

Thus, when comparing nuclear energy, always bear in mind that the sole criterion is the energy per particle.

Let us bring this discussion to an end with one more comparison. How does chemical energy stand up to nuclear energy? The specific binding energy (that is, per particle) in nuclei is roughly a million times the specific chemical energy. That explains why alchemists using purely chemical methods were never able to convert one element into another (actually one nucleus into another, because it is precisely the nuclear composition that determines the structure of the atom and its chemical properties).

Yes, indeed, the energy concentrated in atomic nuclei is stupendous. The exchange of mesons is what cements the nuclear particles together so strongly. This can be put somewhat differently: if (by analogy with an electric charge) we introduce a nuclear charge (also called a meson charge), the latter will far exceed the electric charge.

New facts and fresh conclusions

In our story of the meson interpretation of nuclear transformations we left out a number of very important points that fill out the picture considerably. After Yukawa predicted the new meson particle, experimentalists got busy for the search, which in itself is an exciting chapter in the history of science. Five entirely new particles were finally discovered. Two of them had mass 207 times that of the electron: one was positively charged, the other was negative, and they were called mu-mesons (designated μ^+ and μ^-). For some time they were believed to be the Yukawa mesons. Yet the mu-mesons did not exhibit any activity in interactions with nuclei. At any rate they did not differ at all from electrons in this respect.

Fresh explorations led to the discovery of pi-mesons (π-mesons, or pions), which had all the attributes for the role of carriers of nuclear interaction. There were three kinds of pi-mesons: the positive pi-meson (π^+), the negative pi-meson (π^-) and the neutral pi-meson (π^0). Their masses differ so slightly (273.2 electron masses for the first two and 264.2 electron masses for the latter)

that these particles are justly considered to be the same meson in different states of charge.

Now do they all play some part in nuclear processes? Theory and experiment both say that they do, but the roles are not the same.

Let us start with the pi-plus-meson. It is quite obvious that it cannot be ejected by a neutron, otherwise the neutral neutron would become a particle with negative charge. But there are no such particles.*

On the other hand, the pi-plus-meson is readily ejected by a nuclear proton, which loses its charge in the process and—note this interesting circumstance—turns into a neutron. Now what is the subsequent fate of the ejected pi-plus-meson? It cannot be absorbed by the proton, for the latter would then have a double charge, whereas physicists were well aware of the fact that with respect to all particles the charge can be either zero, or some exact multiple of the electron charge. Consequently, the pi-plus-meson may be absorbed only by a neutron, which takes on the positive charge and then obviously turns into a proton.

So after the ejection of a pi-plus-meson by a proton and its absorption by a neutron, the proton and neutron simply change places, as it were. The situation with the pi-minus-meson is similar, only a neutron ejects the particle and a proton absorbs it. Again the particles change places.

Now since the pi-zero-meson is neutral, it may be ejected and absorbed by any particles irrespective of their charge. Although the pi-zero-meson contributes considerably to interaction, we shall concentrate our attention on the exchange of charged mesons. One is immediately struck by the high degree of symmetry: mesons of both signs work with the same intensity transferring interaction. This is closely associated with what is known as "charge independence" of nuclear forces, which experimentalists are constantly drawing attention to. Charge independence finds expression in

* The antiproton has to be dismissed because if an antiproton appears close to a proton (and there are always protons in the nucleus) they annihilate almost instantly. Nothing like that occurs in nuclei. There are other reasons too, but we shall confine ourselves to these.

the fact that the specific meson (nuclear-proper) forces are identical for any particle pair. There is no difference between proton-proton, proton-neutron and neutron-neutron interactions. Thus, charge independence simply means that meson forces are independent of the electric charge of the particle.

Metabolism in the nucleus

You probably recall that when we were discussing the composition of nuclei at the beginning of this chapter we were confronted by the problem of how stable nuclei could contain unstable neutrons as constituent parts. We have now finally arrived at the answer. Due to the exchange of charged pi-mesons, the neutron turns into a proton so quickly that it does not have time to disintegrate. Something like holding a hot potato—you have to keep juggling it from hand to hand to keep it from falling to the floor. That is how charged pi-mesons are tossed about by nuclear particles; it ultimately produces a stable system. But remember, it is dynamic stability with the intricate internal motions that we have just dealt with.

3

Beta-decay

Is it always possible to establish equilibrium of this kind? Obviously not. It is all right as long as the number of protons and neutrons is the same. But what if there are more neutrons? Then each proton will have to take care of more than one neutron. How far can we go in this direction? When two protons are called upon to

serve not more than three neutrons: that is because the neutron is after all rather stable, and one of them can wait while the other is doing an exchange trick with the proton. But when we go beyond this limit, equilibrium is upset. Sooner or later, one of the neutrons that doesn't get around to interacting with a proton disintegrates. This is called beta-decay (β-decay). The excess neutron decays into a proton (making the proportion of particles in the nucleus stable), and the decay products (the electron and other particles that will be taken up in the next chapter) fly out of the nucleus.

The reverse can also occur: when there are more protons in the nucleus than required. This too leads to instability*, for there may be too few neutrons to handle the protons.

We know that, unlike the neutron, a free proton does not disintegrate. This is understandable because the neutron is heavier than the proton and can turn into the latter by conveying the mass difference to the newly generated particles. What can the proton turn into? First we know that due to the uncertainty relation the mass difference of particles in the nucleus ceases to play a decisive role. The point is that the masses themselves (and also their proportionate energies) cannot at all be specified precisely.

This removes the prohibition from the proton, which can then decay like the neutron under appropriate conditions. Of course there is a difference. The neutron emits a negative particle—the electron, while the proton emits a positive particle—the positron. In positron decay, one of the protons (the extra one) converts into a neutron and restores equilibrium in the particle ratio.

By way of illustration let us take hydrogen. The nucleus consists of a single proton, which, naturally, is stable. If we add one neutron to the proton, the nuclear charge does not change. Which means that the number of electrons in the atom remains the same (unity, in our case) and, hence, there will be no change in the chemical

---

* Important also is the fact that protons repulse electrically though this repulsion is considerably weaker than nuclear (or meson) attraction, the nucleus becomes a more stable system as the proportion of neutrons increases. In other words, there is an energy disadvantage in having large numbers of protons.

properties. The element will not move from its first place in the Periodic Table, but the weight of the atom (actually the nucleus) will be doubled. This goes by the name heavy hydrogen, deuterium. The nucleus of deuterium (often called a deuteron) is quite stable, which it should be, in the light of our foregoing discussion.

Now let us add another neutron to our deuteron. The result is a system made up of one proton and two neutrons (called tritium) and is the nucleus of another hydrogen isotope*. But this isotope cannot be stable: it experiences electron decay (beta-decay) and turns into an isotope of helium with a nucleus containing two protons and one neutron (which just lies within the limits of stability). A system still less stable than tritium is an isotope of hydrogen where the nucleus contains one proton and three neutrons.

We could go on illustrating with other cases, but the principle is so simple that the reader should easily be able to fill in tables of stable and radioactive elements and will start making mistakes only when more subtle and involved factors dealing with particle configuration in the nucleus come into play.

Alpha-decay

We now take up alpha-decay (α-decay). Like beta-decay, it was discovered at the end of last century by Becquerel and was soon thoroughly studied experimentally, first of all by Marie and Pierre Curie, Rutherford and a large number of other scientists. This is what they found. In alpha-decay a particle flies out of the nucleus that carries a positive charge of two (in electron units) and a mass almost exactly four times the proton mass. The alpha-particle has all the features of the helium nucleus, that is, two protons and two neutrons tightly cemented together.

Why does alpha-decay occur? Why is it peculiar only to heavy nuclei? Why do some nuclei disintegrate very

---

* Isotopes are elements with identical chemical properties but with different atomic mass (i.e., with the same number of protons but different quantities of neutrons in their nuclei).

rapidly while others have lifetimes of thousands of millions of years before ejecting an alpha-particle? Those are the first questions that come to mind.

There is a striking difference between beta- and alpha-decay. In beta-decay, particles are ejected from the nucleus that were never there in the first place and which for that reason have to be generated in the very process. In alpha-radiation, the nucleus obviously expels one of its constituent parts.

We would now like to know whether the alpha-particle exists in the nucleus ready-made or whether two protons and two neutrons merge into a unit just before ejection. The current opinion is that the latter mode seems more likely, though occasionally the reverse opinion is also expressed. One thing is definite, however, and that is that the alpha-particle is an exceedingly compact, stable and tightly cemented system that originates inside the nucleus (either some time before ejection or immediately prior to the act).

Now about the forces that throw out alpha-particles. It of course carries an electric charge of the same sign as the nucleus and, hence, there is a repulsive force between the nucleus and the alpha-particle. But inside the nucleus it is heavily outbalanced by nuclear (meson) attraction. This is not surprising because otherwise the nuclei would fly to pieces in no time.

A jump through a wall

But if the forces of attraction exceed those of repulsion, how can disintegration occur at all? This is where quantum effects come in again. Can a potato jump out of a kettle? Of course not, it simply hasn't enough energy. And a pencil cannot break through tightly gripping fingers. The point is that both the pencil and potato are gross, classical, objects (macroscopic is yet another familiar term for them), which consist of enormous numbers of particles. Now the alpha-particle has only two protons and two neutrons. Which means that particle-wave duality and the consequent uncertainty relation should be very pronounced in this case. In a certain sense, the alpha-particle is in a sort of kettle too, the nucleus. Consequently

there should be a certain indeterminacy as to momentum and energy spread. To return to our potato case, it would be as if the potato were constantly maintained in some kind of motion by pushing and bouncing and tossing. Then there would be a realistic possibility that at some instant of time the kinetic energy would increase to a point sufficient for it to jump over the top.

The nucleus differs in many ways of course: there is no wall, no fence. Simply the energy spread makes it possible for a particle to overcome the nuclear attractive forces and fly out of the nucleus. As soon as the alpha-particle gets out, the attractive forces fall off drastically because they are short-range forces. Then electric forces of repulsion take over, which diminish much more slowly. They actually throw the alpha-particles away from the nucleus, accelerating them to high energies. That is why alpha-particles have such enormous velocities.

Anything left out?

There is one obscure link in this chain of reasoning. Why does a whole alpha-particle have to leave the nucleus and not, say, a single proton?*

All the more so since all the arguments for indeterminacy of momentum and energy spread are applicable in this case as well. Something obviously is lacking in our reasoning. Let us take a closer look once again.

In both cases the energy spread is the same per particle. But how about the forces of repulsion? Since alpha-particles have two neutrons, half of the system is chargeless and the repulsion in purely proton terms is double. So if helium nuclei are ejected and not single protons, the obvious reason is that the bonds holding one particle (proton) in the nucleus are stronger than those which hold an alpha-particle.

Saturation of nuclear forces

Let us take a simple model that will help explain the situation. Imagine a set of balls each with four

* Just recently Soviet physicists discovered proton decay in nuclei, but this is an exceedingly rare event

threads attached. We tie them together so that all threads are utilized. One way of connecting them is shown in the accompanying figure (I). This is a case of what might be called equivalent bonds: all balls are in equal conditions. It is easy to see that it is just as hard to jerk loose any group of balls as it is to pull out one (the same number of threads has to be broken).

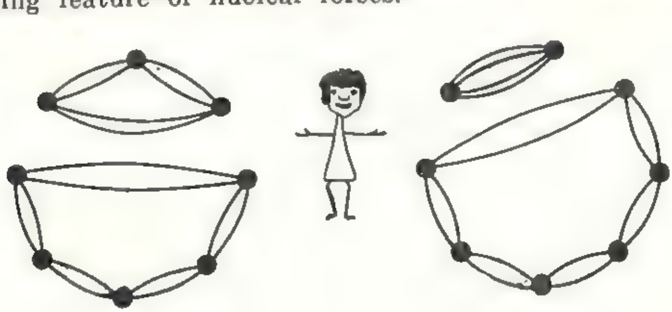
A different mode of connection is shown at II. The situation has changed here: to extract one ball (one particle is on the tip of my tongue), four threads have to be torn. Yet there is a group connected with the remaining system by only two threads. Note that this group is very strong because it has an additional bond (an inner thread). It is precisely this accentuated inner bond that weakens the outer connections of the group with the remaining balls of the system.

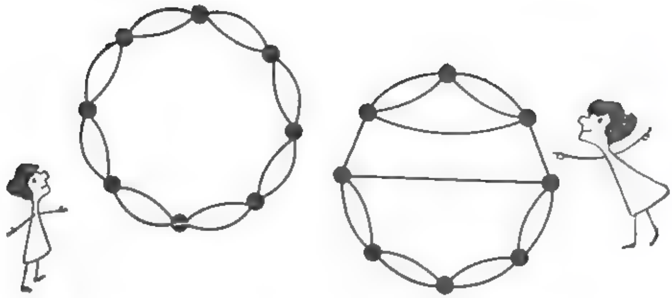
Two more schemes may be devised that bring this feature out still more markedly. Here the strengthened inner connections lead to a total separation of the groups. All the threads are used up on the inner connections so that none is left for external bonds.

And so we come to a situation where the stronger the bonds inside a definite group of balls, the fewer are the threads that connect them with the remaining balls, and, hence, the easier it is for this group to get out of the system.

Nothing like this would happen if we had an unlimited supply of thread. The point is that each ball can link up only with a *limited* number of its neighbours.

This is an exceedingly important circumstance and deserves to be stressed once again. If this is what happens in nuclei, then we have come upon a highly interesting feature of nuclear forces.





But perhaps we are seeking analogies where there are none? Numerous facts, however, convince us that this is a very, very close analogy.

First of all, the alpha-particle is undoubtedly a very tightly bound particle. It is even used as a shell to bombard other nuclei. Perhaps it is precisely this monolithic quality that accounts for the comparatively weak attraction of the remaining nuclear particles.

Now it might be possible (like in our ball model) to picture each proton or neutron actively interacting only with a relatively small number of particles in the immediate vicinity. Are there any grounds for this conclusion? Yes, there definitely are. Actually, we arrived at this conclusion when discussing beta-decay. There, each proton was in a position to accommodate (ensure stability to) only one or two neutrons. But this accommodation involved exchanging charged mesons (with the exception of pi-zero-mesons, which may be considered to establish a uniform "background of attraction" for all particles). However, this very same exchange of mesons is not only a "safety valve" against beta-decay, but plays an exceedingly important extra part as well—it is the vehicle of interaction!

The conclusion is obvious: in interaction events, too, each nuclear particle is able to accommodate only its immediate neighbours. This is because nuclear forces, as we know, operate over very small distances. Physicists explain this in terms of what they call the saturation of nuclear forces. Another interesting fact supporting the saturation notion is the law of approximate constancy of nuclear density. Experimenters have demonstrated that

nuclear dimensions increase as the cube root of the total number of particles in the nucleus. In other words, the volume (which is proportional to the cube of the radius) increases in direct proportion to this number. Hence, the volume per particle is practically the same in all nuclei. Let us try to find an explanation. Suppose that two nuclei merge. We shall go into this in detail a bit later. If all particles then interacted with all the others, a certain shrinkage would have to take place. Due to enhanced attraction, the particles would be pressed in more tightly than before. But this does not happen. The space allotted to each particle does not diminish. Which means that most of the particles (practically all of them with the exception of border-line ones) will not experience any change in interaction. This is because the forces acting on them have reached saturation, and the appearance of fresh particles in the immediate vicinity does not add anything to these forces.

Now that we have learned about saturation and can see that our analogy with the model is justified, the way is open to an understanding of the essentials of alpha-decay.

For instance, heavy nuclei with large numbers of particles are more prone to alpha-decay than others. The obvious explanation is that saturation does not come into full force in nuclei with small numbers of particles.

Now we can answer the question we started out with: Why do groups of four particles (two protons and two neutrons) fly out of nuclei and not separate particles? We can now see that due to a merging of protons and neutrons into an alpha-particle, the bonds linking them with their surroundings are weakened. As soon as that occurs, the energy spread which stems from the uncertainty principle is sufficient to allow alpha-disintegration to occur.

We have thus analyzed two types of nuclear instability. There is one more kind.

Nuclear fission

This is not an ejection by the nucleus of some small group of particles, but a break-up of the nucleus into two more or less equal parts. Only heavy nuclei can fission

that contain over 250 particles. This in itself gives a clue to the process. We know that nuclear forces have very short range, we know about saturation, and it will be easy to imagine what happens in a nucleus with so many particles. Various parts of such nuclei almost live separate lives. Particles on opposite sides of the nucleus are hardly connected in any way. All you have to do is shake up a nucleus like this and it will fall into two pieces. A piece of clay breaks at the slightest touch, even due to its own weight. The same thing happens to a drop of mercury. Note that small droplets of mercury and small pieces of clay are much more stable. That is the picture qualitatively speaking. It appears so simple and clear-cut that we might end our story right here if it weren't for one very essential circumstance. Take a look at the nuclear table (which is practically the same as the Periodic Table of Elements), and you will see immediately that the nuclear mass increases from element to element faster than does the charge. Which means that the number of protons in the nucleus increases more slowly than the number of neutrons.

The reason is quite obvious. There is no saturation in the electric forces of repulsion. Each proton interacts with all the others, no matter how many there are (too, Coulomb forces operate to greater distances). As the number of protons builds up, the forces of repulsion become greater and greater, and can be balanced only by an increase of neutrons in the nucleus because neutrons are insensitive to electric repulsion yet add their bit to nuclear attraction. To build up this attraction faster than electric repulsion, we have to add many more neutrons to the nucleus. Remember that each proton is not attracted by all particles but only by those in the immediate vicinity (don't forget the saturation of nuclear forces). It is quite natural, therefore, that as the number of particles increases in the nucleus, the proportion of neutrons must continue to build up.

Now imagine that some heavy nucleus splits in two. Take the nucleus of the uranium-239 isotope with 92 protons and 147 neutrons. For the sake of simplicity, we picture the nucleus to have divided in half. Then each of the fragments will have 46 protons and 73 or 74 neutrons. The number of protons and, hence, the charge of

the nucleus (in electron units) coincides with the number of the element in the Periodic Table. Consequently, the fragments are nuclei of palladium. But the most stable isotope of palladium has 61 neutrons in its nucleus. Where do the extra 12-13 neutrons per fragment go to? They can of course turn into protons via beta-decay. But since fission occurs very quickly, something more simple will have time to take place (outwardly more simple, naturally). Some of the extra neutrons will be thrown out and become free. This release of neutrons is what makes the familiar chain reaction possible. Indeed, if a certain amount of fissionable nuclei is collected, sooner or later one of them will break up into fragments either due to some external factor or spontaneously. The emerging neutrons will freely fly into neighbouring nuclei (remember they do not experience electric repulsion) and will shake them up sufficiently to fission them. Fresh acts of fission will give rise to more neutrons and the process will avalanche at a fast rate bringing all fissionable nuclei into action.*

Since the fragments in each act of fission acquire tremendous energies (electric repulsion throws them away from each other with a fantastic force), the fissile material releases large quantities of energy, which is carried away by electromagnetic and other radiations, and considerable heat that may be utilized. Millions of people on this earth are fighting to direct this energy into peaceful channels.

Incidentally, the technical application of nuclear fission is somewhat akin to the use of the wheel. Neither operate naturally.

We would stray too far from our story if we were to go into a detailed description of industrial uses of nuclear energy released via the fission of atomic nuclei in nuclear reactors. But since we have started on nuclear sources of energy, a few words might the added about thermonuclear, or fusion, reactions.

---

* This is naturally a rather rough outline of the chain reaction. Actually, not all neutrons are captured by fissionable nuclei, and those that are captured do not all initiate fission.

When nuclei fuse together

Man has always utilized energy released in fusion reactions. That is, since he got acquainted with fire. But this is chemical fusion: the combination of oxygen atomc with the atoms and molecules of the fuel. Why is energy released? Why is there a flame (or an explosion) when hydrogen is mixed with oxygen? Simply because atoms of oxygen and hydrogen have more energy separately than when they are combined into a molecule of water. It is this energy difference that is released in the combustion process.

Similarly, nuclear fuel should burn. The question is how and under what conditions can we make nuclei merge.

At first glance, the simplest thing would be to try neutrons since the only forces operating between them are attractive. But there is no way to store neutrons, they pass through all walls or are absorbed by the wall materials, and we must never forget that they are not stable. The fusion of two protons need hardly be discussed—the electric repulsion is too great. Now deuterium nuclei (heavy hydrogen)—deuterons—are something quite different. They are stable. This is a stable system of one proton and one neutron. Still more stable is the helium nucleus, which we should be able to obtain by bringing two deuterons together. Not deuterons but alpha-particles are ejected in radioactive decay. And what is more, the meson model of interaction leads us, qualitatively, to this same conclusion. In the deuteron, only two particles exchange mesons and there are latent possibilities. If, for example, a proton has thrown out a pi-plus-meson and converted into a neutron, this neutron will have to wait until its partner, having absorbed an emitted pi-plus-meson, turns into a proton. Only then will there be another chance to exchange charged mesons. It is quite obvious that the helium nucleus has more exchange opportunities. But the more active the meson exchange, the stronger the system. Which means that when two deuterons collide they can merge. And energy should be released (quite considerable energy, it may be added), just like energy released in any kind of chemical fusion.

But how about electric repulsion? Yes, it exists. What is more, it complicates matters greatly when we try to bring two deuterons together. But if the resistance of the electrical forces is overcome and the deuterons are close enough for the short-range nuclear attractive forces to take over, they will suppress repulsion completely.

The most important thing therefore is to bring the deuterons to a close distance, and the work done will be covered a hundred-fold. So the problem is to bring the deuterons together.

One method is to heat heavy hydrogen to temperatures measured in tens of millions of degrees. At these temperatures, the energy of thermal motion becomes sufficient to overcome the armour of electric forces. In collisions, nuclei come so close that meson interactions can hold them. A thermonuclear (fusion) reaction then ensues. This is a merging that occurs at ultrahigh temperatures. In the process, enormous quantities of energy are released; this boosts the temperature still more. The process of nuclear "burning" becomes self-supporting as long as there is fuel available.

In nature, fusion reactions are not at all rare. Thermonuclear fusion keeps all the stars burning, including our sun, and is the ultimate source of all life on earth. Stellar fusion is somewhat more complicated, though, and is not simply a formation of helium out of deuterons. A whole chain of nuclei is involved, yet fundamentally the essentials are the same.

Only a few more words need be added to complete our story. The fusion reaction involves mainly light nuclei (the more protons in the nucleus, the harder it is to overcome electric repulsion). But there is still another very important circumstance. In extremely heavy nuclei, saturation of nuclear forces becomes pronounced and then fusion is impossible. Here, the role of saturation of nuclear forces is somewhat reminiscent of that of the saturation of chemical forces in the formation of molecules which was mentioned earlier in our story. Thus the choice of fuel is rather narrow.

We have not yet been able to accomplish a controlled fusion reaction, but scientists are hard at work in numerous centres of research. The great effort is justified by a truly

great goal, because harnessing the thermonuclear reaction will mean a practically limitless supply of energy for humanity.

What have we learned?

We can now bring to a close our story of nuclear forces. We have learned a great deal of interesting and important facts. In a literal sense we cannot speak about forces in the atomic nucleus because force is a purely classical and nonquantum quantity equal to the product of the mass by the acceleration. Wave-particle duality does not admit of an exact specification of coordinates and velocity (and, hence, acceleration). Consequently, we cannot speak of any kind of forces in the microworld in the mechanical sense. Here there are other criteria of interaction. The simplest is the mean binding energy. Recall the uncertainty relations for energy and time. Stable nuclei have, practically speaking, indefinitely long lifetimes. For them the indeterminacy of time may be considered infinitely large. But then the indeterminacy of energy (note that it is for the whole nucleus and not for the constituent parts) has to be infinitely small, since these indeterminacies are reciprocals. It is precisely the absence of an energy spread that enables us to retain energy as a characteristic of interaction for a purely quantum entity like the atomic nucleus.

We have shown that interaction is due to an exchange of mediating particles. Evaluating the mass of the latter, we obtained a meson pattern of interaction. In doing so we were able not only to give a spectacular explanation straight off of important intranuclear interactions (their short-range nature, saturation, so-called charge independence, which states that nuclear forces act on a particle irrespective of whether it carries an electric charge or not), but also to resolve one of the most fundamental paradoxes: the stability of nuclei that are built up out of such unstable particles as neutrons.

We even went further. We explained the basic features of beta- and alpha-decay, deciphered fission reactions and the fusion of atomic nuclei. Using the results obtained, we can, for instance, point out that pi-mesons are not the only carriers of nuclear interaction, though they play the

main role. Any quanta capable of being emitted and absorbed by nuclear particles can be the vehicles of interaction. And the heavier these quanta, the smaller the range of the respective forces. To illustrate, we can take the so-called *K*-mesons discovered quite recently. They have roughly 970 electron masses (this is more than threefold the mass of a pi-meson) and, therefore, the interactions effected by these particles should extend to distances only one-third of the pi-meson distances.*

There is another exceedingly important problem that immediately comes to the fore in any discussion of nuclear forces. A proton emits a pi-meson, which is absorbed by a neighbouring neutron. But why only a neighbouring one? The proton turns into a neutron in the act, and the new neutron, being no worse than any of the other ones round about, can capture its "own" meson. A neutron can indulge in a similar exercise**. As a result, an interaction is established of a particle with itself and not with other particles, something like electromagnetic self-action. The important thing here is that a proton and neutron should appear as a very intricate system, in the centre of which is a sort of core surrounded by a cloud of mesons undergoing constant emission and absorption. (Incidentally, we might add that the interaction of different particles may be pictured as a partial overlapping of such clouds.) Mesons are charged particles, so we can raise the question of the distribution of electric charge in this cloud. Which is a definite step forward in determining the *structure* of elementary particles! Just yesterday the word "elementary" was for many synonymous with "structureless". This talk about the structure of particles is not just theoretical conjecturing. The structure was actually probed experimentally in Hoffstadter's elegant experiments that we mentioned earlier in this book. A remarkable confirmation of the correctness of the theory!

---

* The role of *K*-meson interaction is particularly significant in so-called hypernuclei, i.e., nuclei made up of protons, neutrons and hyperons—extra-heavy particles with masses of about 2,200, 2,300 and 2,500 electron masses.

** There is obviously no sharp distinction between a proton and a neutron. They are more likely different states of the same particle (charge states, so physicists say) with the generic name nucleon.

Blank spots

Then where do the blank spots come in that we mentioned at the start of the chapter? *Such tremendous advances made in both theory and experiment. Nuclear interactions explained, experimental probing into the interior of* particles themselves! True, very true, but all this has been done only qualitatively, descriptively, but not at all quantitatively!

As soon as the physicist attempts to translate his discussions into the rigorous language of equations and formulas, a whole forest of difficulties arises, many of which (actually, almost all of them) have still to be surmounted. And there are points, too, where theory cannot even give a qualitative description. For a number of reasons, physicists had to introduce the notion of new forces (forces of repulsion) operating over exceedingly small distances ("of the order of the dimensions of the particles themselves" is the way they put it). What kind of forces are they? Opinions differ, but frankly speaking no one knows for sure.

We are still in the dark about many other, still simpler things. We are not very confident about the shapes of the various nuclei and particle configurations within them.

This is the region of blank spots that is being attacked by a whole army of scientists.

We already know that the qualitative picture described above is essentially correct. Quantitative theories are appearing. Some are based on familiar quantum-field conceptions, others have tried new avenues of approach based on fresh hypotheses.

One can be sure that each important step in the physics of nuclear forces will help answer the basic question of physics: How is matter constructed?

CHAPTER FIVE

Deep thoughts are there
That bridge the gap faithlessly
Between creation and decay
From atom to star.

(Emile Verhaeren)

WEAK INTERACTIONS

1 *The Disintegration of Elementary Particles and the Neutrino*

2 *The Charge and the Transformations of Elementary Particles*

3 *The Neutrino and the Evolution of the Universe*

1

The hunting grounds of science fiction

Science-fiction writers are constantly having their heroes turn forces off and on. We find gravitational screens that destroy the effects of gravity, radiation that breaks up chemical bonds, devices that circumvent friction, and what not.

But we haven't found anyone who has tried to picture the world without weak interactions.

Yet this is a whole fantasy world in itself.

Weak interactions are also called "decay" interactions. They involve the disintegration of almost all unstable particles. We mentioned this when discussing the mutual transformations of particles.

If these interactions suddenly disappeared, we would be lacking some very common types of particle transformation. Neutrons, many mesons and hyperons would become stable with indefinitely long lifetimes.

Wonderful things would come about at every turn. Take the Periodic Table of Elements. Today it has 104 squares, 104 chemical elements recorded by science.

And why not more? Are there elements with numbers 200, 1,000, or more?

No, there are not. And what is more, we are positive that they will never appear in the Periodic Table. The reason is probably obvious to the reader who has gone carefully through the preceding chapter.

The point is that the number of an element coincides with the quantity of protons in the nucleus. The more there are, the greater the Coulomb forces that strive to break up the nucleus. The only balancing force is the considerable neutron population that adds nothing to the forces of repulsion and cements the nucleus via the forces of meson attraction.

It would seem that all we needed to do was dilute the protons with a sufficient number of neutrons and thus counteract the Coulomb instability of any nucleus. But here we must recall the instability of the neutron. As soon as the number increases substantially, there is a possibility of beta-decay, which increases as the relative proportion of neutrons in the nucleus builds up.

Thus, superheavy nuclei cannot be stable. This well-known circumstance explains why the very heaviest elements are not, strictly speaking, discovered, but rather "manufactured". These elements are not found in the pure form either in the interior of the earth, in the atmosphere, or deep under the water, for they are too short-lived. Scientists make them by firing high-speed particles and keeping tabs on a whole series of complicated nuclear transformations until ultra-sensitive instru-

ments record, in a brief instant, the new element, which comes to life in such minute quantities that the individual atoms can be counted.

Now what if neutrons did not decay and were stable? Then, of course, nothing would prevent us from building up larger and larger quantities of them. The Periodic Table would continue to grow; true, not indefinitely as might appear at first glance. Recall the saturation of nuclear forces. Giant nuclei would be very unstable and would easily fall apart. True, to some extent we could prevent fission by taking special precautions against any jolting or shaking. We might picture a sign on the laboratory door: "Caution! Weak interactions off." On the table, under an opaque hood (to eliminate any shake-up due to light rays) an amorphous body the size of an apple floats in liquid helium (the temperature must be as low as possible to reduce thermal jolts). Incidentally, there would be no floating because the body would be extremely heavy, weighing nearly a million tons. Why it wouldn't break up under its own weight would indeed be a mistery because a crack of even a thousand millionth of a millimetre across would disrupt the short-range nuclear bonds, and fantastic forces of electrostatic repulsion would throw out fragments of this hypothetical element at tremendous velocities.

On the hood is inscribed: "Element No. 1,000,000,—000,000,000,000,000,000,000,000" and a few words on life insurance.

A truly fantastic picture even in a fantastic world without weak interactions!

Many things have been "forgotten". If there were no weak interactions, and if the neutron were a stable particle, the table of elements would continue for many tens, even hundreds of numbers. But still more amazing things would happen to the table of isotopes.

In the chapter on nuclear forces we said that isotopes have the same number of protons in the nucleus but different numbers of neutrons. An isotope is stable only when the ratio of protons to neutrons lies within a "stable limit". As soon as the number of neutrons goes beyond this limit, we have beta-decay. If it weren't for the weak interaction, there would be nothing to fear about beta-decay and the possibilities of increasing the number of

neutrons in the nucleus would be extended enormously. Hydrogen would then have not four isotopes (of which ordinary hydrogen and deuterium are stable) but a practically indefinite number. True, some place around the thousandth isotope, a new cause for instability would arise: atomic electrons would begin to come in contact with the overgrown nucleus round which they revolve, and what is more, a thermal instability, which we have already mentioned would set in. But these are only incidental circumstances relative to the intranuclear situation.

In our conjectured world, there could be still another type of stable exotic nucleus. In the Periodic Table, it would be situated in a zero square, under hydrogen. This is a nucleus without any protons. If neutrons did not disintegrate, there could be hundreds or even millions and more neutrons existing as stable systems. And they could be regarded as isotope nuclei of this bizarre element that has no atoms in the ordinary sense of the word (neutrons do not attract electrons). Atoms without electrons, devoid of chemical properties!

This kind of atom would not be inert, however, as might be thought due to the fact that it has no chemical properties. It would be a very dangerous neighbour to have about, for whenever it found itself in the vicinity of another atom, this bundle of neutrons would pass through the electron cloud without hindrance and would be attracted to the other nucleus with a tremendous force. A typical nuclear fusion would occur, but it would not be a thermonuclear reaction because no heating at all would be necessary. If the bundle of neutrons were large enough, a perceptible explosion would result.

So you see, even the "pure nuclear substance" of science fiction is aggressive.

But why confine ourselves to neutrons? The word instable can be applied to nearly every square in the table of elementary particles. Instability, as we have pointed out a number of times, is, with few exceptions, associated with what we tentatively call the weak interaction. If it weren't for the latter, then not only neutrons, but also mu-mesons, charged pi-mesons, *K*-mesons, and also particles heavier than protons and neutrons (called hyperons) would be stable. For instance, the mu-mesons. In

many respects they are very much like electrons and positrons. They are charged both negatively and positively. But that is not all. The similarity is so great, in fact, that physicists at times believe that the negative mu-meson is actually an electron that has somehow gained in weight. The mu-meson is 207 electron masses.

But how about decay? Isn't that an essential difference? The electron is stable, while the mu-meson lives only millionths of a second. For an answer take the following example. Imagine an atom in the excited state. An excited atom is unstable. As a rule, it decays almost instantly to an unexcited atom and photon. Yet we do not say that an excited atom and an unexcited atom are different systems. We prefer to use the expression: one and the same system in different states. Perhaps the mu-meson is an excited electron.

We have strayed away from our topic, however. We know how important is the work electrons do. They form the cloud of the atom, and hence determine the chemical properties. Electron motion is the cause of electric current in metals; the electron is the main participant in all kinds of cathode-ray devices, from the simplest diode (a two-electrode tube used in electric-current rectifiers) to electron microscopes and betatrons. Electrons play the leading role in present-day science and technology. But couldn't mu-mesons do the same? Instability is the disturbing factor. If it weren't for instability, mu-mesons could take on all the electron functions and do a good job, sometimes even better than the electron.

So far we have been going on the premise that there are no weak interactions. True, atoms have actually been found with mu-mesons (negative, naturally) taking the place of electrons. Though these mu-mesonic atoms live only a short time, scientists have recorded their entire spectrum. This is extremely interesting in view of the fact that the meson orbits are 207 times closer to the nucleus than electron orbits (since a meson is 207 electron masses). For this reason, the meson is much more responsive to all the peculiarities of nuclear structure and informs us about them via its spectrum.

Since we have already started on systems that include mu-mesons, it is worth mentioning another curious possibility. Imagine something like a hydrogen atom

with a positive mu-meson as nucleus. If the mu-meson were stable, such atoms could be built into molecules. Then we could obtain unusual chemical compounds like "extra-light water".

A world without weak interactions is really fantastic. Let us wind up our science fiction with a bit about hyperons. If hyperons were stable, we could enrich our set of atomic nuclei enormously. We could produce stable nuclei made up of mixtures of neutrons, protons and various hyperons, and even make nuclei solely out of hyperons. Using neutron hyperons, we could even build electrically neutral pieces of hyper-nuclear matter.

By this time the reader has definitely got a picture of how important weak interactions are after all. They involve so many prohibitions (and simplifications, some would think). But the term "weak" does not in the least mean insignificant.

However, there is good reason for this term, as we shall soon see.

Elusive for a quarter of a century

The word "neutrino" appeared about thirty years ago. The name (meaning "small neutron") was given to a new particle, probably the most remarkable and popular in the family of elementary particles.

It made its entrance into science in a strange way; its properties are fantastic, and its role in nature is extraordinary.

This particle was hypothesized so as to keep the foundation of physics intact, it was needed to save the conservation laws. Proof of its existence was first obtained experimentally only in 1956. For about a quarter of a century, it existed elusively on the pages of scientific journals and books. Though it had never been "seen", this particle played an important part in the interactions of many particles. First among them (in chronological order) was the neutron.

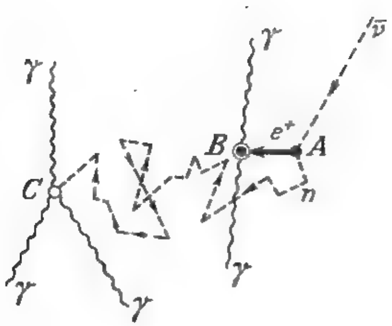
We have already mentioned the beta-decay of the neutron several times. The proton and electron generated in this disintegration process are readily detected by existing apparatus. But the strange thing is that there is

a discrepancy between the energy of the neutron prior to decay and the energy obtained by the proton and electron formed from this neutron. A portion of the energy seems to vanish! Also, there is a paradoxical nonconservation of momentum and angular momentum.

The conservation laws are the most fundamental principles that physicists have been able to establish on the basis of numberless experiments and their interpretation. Specific methods of describing motion may change—the Newtonian description was superseded by a quantum-mechanical description—but the laws of conservation have always remained unchanged. What is more, they have been the beacon that has helped scientists to probe the unknown.

And then beta-decay upset the apple-cart. Physicists were worried. Opinions divided. Some thought that the conservation laws no longer held. They pointed out that these laws were established for the world of "large things", the macroworld, and that for elementary particles they would hold only on the average. This approach did not remove any of the problems, nor did it appeal to most physicists, in view of the fact that it did not contain a positive program for further advances.

Much more attractive was the hypothesis advanced by the Swiss theoretician Wolfgang Pauli. He supposed that a third particle (in addition to the proton and electron) is born when a neutron decays. This particle was



thought to carry off the energy, the momentum and the angular momentum. We do not observe this particle, which is easy enough to explain. It has no electric charge and its rest mass is very small or even zero. Which means that it cannot tear electrons out of atoms, break up nuclei or do any of the "damage" that ordinarily enables us to judge the existence of particles.

We cannot, of course, say that such a particle does not interact in any way at all. What was born can be absorbed. Otherwise, the invention of the neutrino would signify another rejection of the conservation laws, only in a more subtle form, because energy would be lost irretrievably and without a trace along with the neutrino particle.

Pauli presumed that the neutrino exhibits very weak interaction with matter and therefore can pass right through a big thickness of matter without displaying itself in any way. We now know that Pauli was quite right in this respect. The neutrino is indeed an elusive particle. It can pass right through the earth, even the sun, without hitting anything. Only if we conjure up a massive chunk of iron the size of our whole Galaxy, would the neutrino definitely be absorbed.

The godfather of the neutrino and the one who gave it a name was the celebrated Italian physicist Fermi, who legalized the particle by introducing it into the framework of quantum theory.

Fermi's investigations and those of his many followers would seem to have clarified the situation. The rest mass of the neutrino proved to be zero, like that of the photon or light particle. This simply means that there

are no neutrinos at rest. As soon as they are born they begin to move with the velocity of light. The spin of the neutrino is known too. It is the same as that of the proton or the electron.

Information about the neutrino continued to accumulate. Theoreticians predicted that it should have a mirror image, like the electron has its positron. This was then called an antineutrino. It is an oddity that the particles formed in neutron beta-decay should, for a variety of reasons, be called antineutrinos instead of neutrinos.

Experimenters have collected a lot of information about particle transformations in which neutrinos and antineutrinos participate. The list of such transformations (which we shall discuss a bit later) is already rather long. It turns out that neutron beta-decay is by far not the only process that involves these invisible particles. How to catch them was a big problem. But experimenters solved that as well. The experiment was simple in the extreme. A massive "box" was positioned next to a nuclear reactor which produced enormous numbers antineutrinos via beta-decay. The walls of the box were made of such material (lead and paraffin) and of such thickness that no particles could get through. None but the antineutrino. Which is understandable, since practically no barriers exist for these particles. Streams of antineutrinos move out in all directions and of course towards the box. These fluxes are so great that although a single antineutrino has a negligible probability of being absorbed in the matter in the box, several acts of absorption should occur within a relatively short time. Calculations indicated that the process should occur as follows. Let an antineutrino ($\bar{\nu}$) collide with a proton at A (the box was filled with water) turning it into a neutron and producing a positron at the same time. The positron is straightway annihilated when it encounters an electron (at point B), yielding two gamma quanta, which pass through a layer of liquid scintillator (a substance that scintillates when gamma quanta pass through it) situated near the inner walls of the box. This scintillation is immediately recorded by 150 photomultipliers (instruments that respond to minute flashes of light). What about the newly formed neutron? After a short time of wandering about in the water it was expected to be captured by a piece of cadmium in the box (at C), which is

again accompanied by the formation of gamma quanta. As you see, a large number of events were to accompany the capture of an antineutrino. Such were the predictions. But what did the instruments have to say? Did they record everything that was predicted?

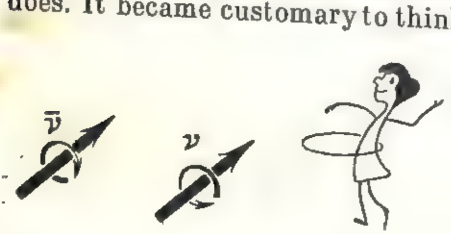
They certainly did. The invisible antineutrino was caught at last. Physicists appeared to have things under control as far as the neutrino and antineutrino were concerned. Theoreticians had described them confidently and experimentalists were able to detect them. But researchers were in for yet another surprise. Nature reminded them to be wary when dealing with the neutrino.

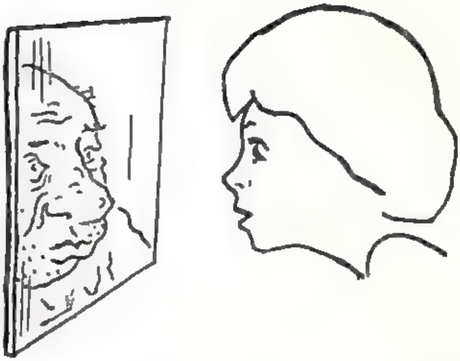
Neutrino the saviour, and neutrino the destroyer

The birth of the neutrino saved the important conservation laws. But this very same neutrino destroyed another extremely general law. Up until 1956 it had never entered anyone's head to doubt the mirror symmetry of nature, which meant that for any process occurring in nature, its mirror image could also occur. In other words, the mirror image of any entity is also a possible object of nature. True, a person examining himself in the mirror might, after some thought, note certain curious details: the right is turned into the left. The mirror image writes with his left hand, though there are lefties; he buttons up his coat on the left side, though it is only custom that makes us do otherwise; the heart is on the right-hand side, though, as we know, there are certain rare individuals with this anomaly. In short, there are no wonders behind the looking glass that could not exist in front of it.

There does exist a mirror symmetry, a symmetry of right and left. But does it exist at all times?

For a long time there was no reason to doubt that it does. It became customary to think so, and what is custom-





ary begins to become indisputable. Neutrino studies, however, once again reminded physicists that science does not countenance truths taken for granted.

We have already mentioned the fact that the neutrino has spin, which is its angular momentum proper. In "classical" language, one would say that it appears to be twisted (like a bullet coming out of the barrel of a gun). Neutrinos generated in the decay of an antineutron are twisted in a very specific fashion: their rotation is that of a left-hand screw in the direction of motion. There are no exceptions (say, like people born with the heart on the right-hand side). But this is an obvious violation of mirror symmetry because a screw with a left-hand thread will, in a mirror, appear to have a right-hand thread. But there is no such thing as a right-hand-screw neutrino. The neutrino is the only particle that does not have a mirror image.

Of course, this does not mean that if we put a neutrino in front of a mirror (which we could do only in our imagination anyway), we would not see its reflection. The point is that this reflection (if reflection we can call it) should have properties that no neutrino is ever able to have under any kind of conditions. The remarkable thing, however, is that these properties are just the same as those of the antineutrino.

Thus, the mirror reflection of a neutrino is a different particle—the antineutrino. Almost as if the mirror image of a beautiful young girl were that of an old bald man.

The main thing is not how attractive the neutrino or antineutrino is but the fact that they are different parti-

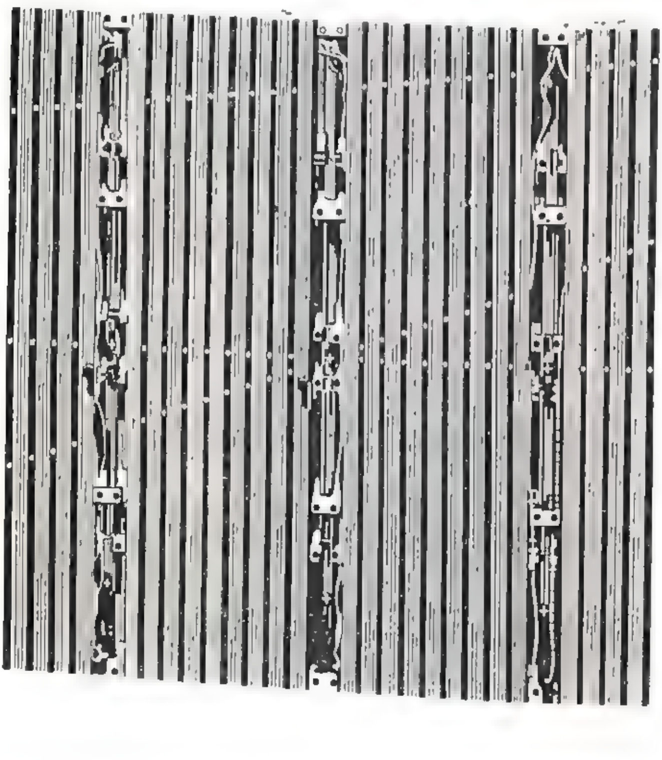
cles, different yet mirror-symmetrical. The establishment of this fact meant the downfall of the simple symmetry of left and right, which was a great surprise to physicists.

Two sorts of neutrino

But that wasn't the end of the surprises. A new marvellous event in neutrino physics took place in 1962.

We have already spoken of mu-mesons. Their similarity to electrons (and positrons if we have in view positive mesons, μ^+) extends also to interactions with neutrinos.

Experiments conducted in 1956 demonstrated that positrons are generated in collisions of antineutrinos with protons. The question is: Why not positive mu-mesons? To that, physicists replied that there just wasn't enough energy. Positive mu-mesons are roughly 200 times heavier than positrons and, consequently, they require that much more energy for their generation. Now an antineu-



trino ejected from a nuclear reactor does not possess such a reserve of energy. But what if it did? Then, say the neutrino physicists, positive mu-mesons would be born just as frequently as positrons.

Now if someone insisted that fast antineutrinos coming out of a reactor generate positrons, many physicists would probably be skeptical, for if that were the case, then we would have to recognize some difference between an electron-type neutrino and a muon-type neutrino. This would mean that there are different types of neutrino which doesn't fit into the accepted neutrino concept. Even such a young science as neutrino physics has developed customary notions and a tradition.

The problem of two neutrinos became topical only when there was an actual opportunity to resolve it experimentally. The idea of an experiment was suggested by the Soviet physicist B. Pontecorvo. And the experiment itself was brilliantly executed by American scientists.

Neutrons are a very convenient source of antineutrinos. However, to produce high-energy antineutrinos one has first to impart considerable energies to the neutrinos. But there are no neutron accelerators. These particles are neutral and we are able to accelerate only charged particles.

But there is another way. We know that a pi-meson decays into a mu-meson and a neutrino (or antineutrino). But what kind of neutrino—the electron type or the muon type? A short time ago the question was never asked. Now that it is asked we can give a cautious answer: at any rate, it is definitely a muon-type neutrino. It is closely bound up with the mu-meson by its common birth. But experimentation is required to determine whether it is also at the same time an electron-type neutrino.

The experiment conducted in 1962 on the 30,000-million electron-volt accelerator at Brookhaven, U.S.A., after two years of preparation was something like this. A beam of accelerated protons bombarded a beryllium target generating streams of pi-mesons, which in turn decayed into mu-mesons and, what is most important, into antineutrinos (and neutrinos) of high energy. True, there were very few of them compared with reactor experiments. However, calculations showed that fast antineutrinos interact more readily with other particles than do slow

antineutrinos.* A spark chamber was used to register the particles generated by antineutrinos. This chamber contained 10 tons of aluminium plates between which a high voltage was set up. When a fast charged particle passes through a plate, a spark discharge occurs between the plates in the gaps along its path of travel. The fiery trace is clearly visible on a photograph and makes it easy to distinguish mu-mesons from positrons and electrons. Special shielding was required so that only neutrinos (and antineutrinos) could get into the chamber from without.

Observations continued six months. During this time, there were only fifty cases of particle generation (don't forget that these are weak interactions). And without exception they were all mu-mesons! Not a single electron or positron!

This was amazing! The existence of two different types of neutrino (and antineutrino)—electron-type and muon-type—was proven.

But what are these types? What difference is there between them? What laws govern them and their actions? So far we do not know. This is a new mystery that science will have to resolve.

The impression of the neutrino is that of a rather unruly and ungrateful particle. True, there was a time when it helped physicists and saved the law of conservation of energy, but since then it has brought forth riddle after riddle. The neutrino has been under study for over three decades, yet we must begin from the beginning again.

Well, not exactly. Elusive as this particle may be, we have succeeded a number of times in bringing it out of hiding. At present we know quite a bit about it and a lot more can be conjectured.

For instance, we know a good deal about the interactions of neutrinos with other particles, about the decays that the neutrino participates in, and about the transformations that the neutrino gives rise to.

Neutrino alchemy

By way of illustration, here are some disintegrations in which the neutrino appears in the first generation (the

* "Fast" and "slow" only signify energy differences because the speed is always the same, that of light.

decay products can themselves be unstable and disintegrate into neutrinos. *)

$\mu^- \rightarrow e^- + \nu + \bar{\nu}$ is the decay of a negative mu-meson into an electron, a neutrino and an antineutrino.

$\pi^+ \rightarrow \mu^+ + \nu$ is the decay of a positive pi-meson into a positive mu-meson and a neutrino.

$$K^+ \rightarrow \mu^+ + \nu$$

or

$$K^+ \rightarrow \mu^+ + \nu + \pi^0 \text{ and } K^+ \rightarrow e^+ + \nu + \pi^0.$$

These three decay patterns of the positive K-meson are possible because the K-meson is a comparatively heavy particle. The reserve of mass here is sufficient to produce three fragments. When the neutral pi-meson does not form, the excess energy is divided between a mu-meson (or positron) and a neutrino. Finally, recall the familiar case of neutron beta-decay into a proton, an electron, and an antineutrino.

These reactions (and others that we have not mentioned) have some remarkable peculiarities. First of all, the symbols that designate the particles can be transferred from one side of the arrow to the other (true, in that case, particles have to be changed to antiparticles).

What is more, the sense of the arrow can be changed as well. This means that each reaction can proceed in both directions.

Let us try this with the beta-decay reaction of a neutron. The first time we wrote it down as follows:

$$n \rightarrow p + e^- + \bar{\nu}.$$

Now let us transfer the electron to the left and change the direction of the arrow. We then get the following reaction:

$$n + e^+ \leftarrow p + \bar{\nu}.$$

But this is a familiar scheme, indeed—the reaction in which the antineutrino was discovered! We can read it this way: the collision of an antineutrino and a proton

* For the meaning of symbols, go back to Chapter Four.

gives rise to a system made up of a neutron and a positron.

Similar juggling of symbols enables us to predict a whole series of particle reactions.

Let us return once again to the problem of two neutrinos. We consider the decay reaction of a pi-meson, for example, the positive pi-meson:

$$\pi^+ \to \mu^+ + \nu.$$

Strictly speaking, the symbol ν does not suffice any longer to denote the neutrino. Insofar as this particle appears in the company of a mu-meson, it is natural to call it a muon-type neutrino and designate it ν_μ.

Now let us recall our rule. It will enable us to write down an interesting reaction:

$$\nu_\mu + \pi^- \to \mu^-.$$

This signifies that when a muon-type neutrino collides with negative pi-mesons (which are always found in ample quantities in the meson cloud around any proton or neutron in an atomic nucleus) it should generate mu-mesons and not electrons.

Such is the reasoning that forms the basis of the theory for experiments to detect two neutrinos.

There is no use in writing out all the neutrino reactions, and we don't intend to do so. It is more important to determine what is meant by "weak interaction".

2

Once again: what is charge?

We have already mentioned the change that the concept "charge" has undergone. Let us recapitulate in brief what we already know. It must be stressed that what we are

going to talk about now is the real essence of "charges" as we at present understand them and, hence, the essence of the interactions they give rise to.

The electric charge is the oldest in the family (if we disregard the gravitational charge, which occupies a rather special place). Its childhood is connected with classical, nonquantum theory. What is more, simply with mechanics. And mechanics, as you recall, is built on the basis of descriptions in terms of forces. No wonder then that for a long time the electric charge was regarded simply as a measure of the action of one charged material point on another.

The Maxwellian approach to electromagnetism did not change much in this respect. The accent was on the mediating element of electrical and magnetic interactions, the field. The charge, as before, remained a measure of the force with which the field operates on bodies. True, this did not exhaust its functions. According to Maxwell, this same charge is also a measure of the ability of bodies to create a field.

Quantum ideas introduced new details. The "force" description was of no importance any longer. The interaction of electrically charged bodies appeared as the result of an exchange of quanta of the electromagnetic field (photons). If we did not know that there are no supplies of photons inside electrons, we could imagine that electromagnetic quanta are capable of flowing into and out of particles, like a liquid through holes. Then the charge would specify the width of such openings. The wider they are, the greater the flux of quanta. But there are no such supplies, and we simply say that the electric charge determines the intensity of emission (or absorption) of photons by charged particles or groups of them.

There is one important circumstance here that has been in the shade in all preceding chapters (perhaps because the approach was somewhat different then).

Here it is. Any particle, whether electron, proton, charged pi- or mu-meson, or any other one (take your pick of any in the table that has an electric charge) that emits or absorbs a photon does not experience any transformations. To be more exact, we must say "almost none" because these particles lose or acquire energy in the act. But this only has to do with changes in the state of motion.

To summarize, then, interaction with photons changes the state of motion of a particle, but does not give rise to mutual transformations (with the exception of decay processes of a neutral pi-meson and a neutral sigma hyperon into two gamma quanta). This marvellous circumstance enables electromagnetic interactions to appear often in a nonquantum apparel.

Nuclear interactions do not have this property or at best possess only some remnants of it. By analogy, we might say that a money economy resembles a natural economy: we have the same "exchange", only in the former case we always use certain quantities of the same monetary units.

In nuclear interactions we also have an exchange of quanta, only these are not photons but pi-mesons. We can again speak of a charge—the nuclear charge, in this case—as a measure of the intensity of emission (by protons and neutrons) of quanta of the pi-meson field, which conveys the interaction.

There is an essential difference, however. The emission of charged pi-mesons involves a transformation of particles (or sources). This is quite different from electromagnetic interaction.

We may say that the process of nuclear (or strong) interaction is accompanied, generally speaking, by a mutual transformation of the particles. There is only one case when this does not occur: when a neutral pi-meson is emitted or absorbed. In all other cases, nuclear interactions involve not only a change in the state of motion, but also a change in the type of particle.

In this light too, the nuclear charge appears as a quantitative measure of the intensity of the mutual transmutations.

Still and all, this aspect of the process is not very marked. The proton and the neutron have a great deal in common. If electromagnetic interactions should suddenly be switched off, these particles would be hardly distinguishable. That is why it is often said that they are not different particles but different charge states of one and the same particle. But nucleons are not the only particles typified by the strong interaction. Hyperons (which together with protons and neutrons fall into a common group of heavy particles called baryons) and also

K-mesons operate via the strong interaction. Hyperons are particles heavier than the neutron and proton and are designated as Λ^-, Σ^- and Ξ—particles. Here we encounter the specific feature of the strong interaction—it involves the transformation of one particle into another.

Finally, the weak interaction. Physicists rarely associate it with anything even remotely related to the action of a force. And the charge here is, as a rule, simply the "constant of the weak interaction", thus again stressing that it is far removed from its classical analogue.

Yet the constant of the weak interaction has every right to stand alongside the other charges of quantum theory, because in this theory the charge determines the speed with which particles convert into other particles: the electric charge in the transformations of charged particles into the same particles (but with an altered state of motion) plus photons; the nuclear charge in the mutual transformations of baryons with the participation of pi- and K-mesons; and, finally, the weak charge, the constant of the weak interaction, like the other charges, describes the intensity with which transformations occur with the participation of the neutrino (and antineutrino). Later on we shall see that this does not exhaust the role of the neutrino.

Perhaps the most important conclusion of this book

To summarize then: any charge determines the activity of transformations in specific groups of related processes. (Actually, the word "related" often simply signifies that only one interaction constant is involved in a given class of transformations.) Today, there are four such charges, if we include the gravitational charge. Only four! And the entire totality and multiplicity of events in the world about us reduce to just these four!

One of the most important problems of physics today is to measure the different charges as precisely as possible, which is not always so easy as it seems. The days of Coulomb and even Millikan have passed.

A more natural thing to do is to compare the energies of the various particle interactions (for specified distances

apart) and not their charges. If we agree to accept the energy of nuclear (meson) interaction as unity, then electromagnetic interactions will come out to 10^{-2}, and the weak interactions will be 10^{-14} of this magnitude.

Quite obviously, the term "weak" is entirely justified for these interactions. But never forget that "weak" has nothing to do with the idea of "insignificant".

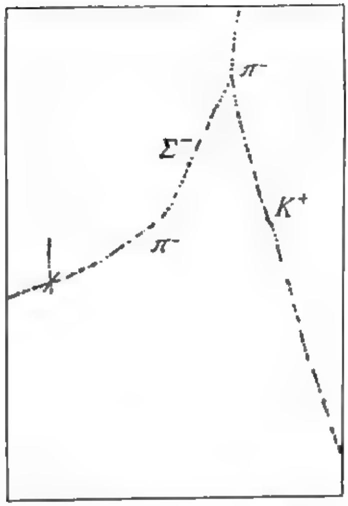
Strangeness enters physics

In a marvellous popular review of the state of the physics of elementary particles, published by Gell-Mann and Rosenbaum in 1957, the authors took for an epigraph the words of Francis Bacon to the effect that perfect beauty is never devoid of some portion of strangeness. Anyone closely connected with the ideas of modern physics cannot but feel, among other things, an aesthetic perfection. One of the greatest of modern theoreticians, Dirac, once said that a physical theory must be mathematically elegant. At the same time the theory is always imbued with a certain strangeness. Even the word "strangeness" has found an application in physical terminology. This has nothing to do with radiant hopes or a sense of humour of the young physicists that put it in circulation. Nature continues to teach lessons to and surprise those theorists who calm down too soon.

Since 1947, the table of elementary particles has acquired a whole bevy of hyperons and *K*-mesons. They appeared all of a sudden and did not in the least follow from any theory. They were "strange particles". And that is the name they were given.

The new particles were not long in justifying their designation. They appeared all of a sudden, and their very appearance took place in an unusual manner. These particles are never generated one at a time, but always in pairs or in greater numbers. Nothing strange in that, one would think.

We know any number of cases of pair production: electron and positron when a gamma quantum collides with a nucleus, and the generation of other particles together with their antiparticles. But the point is that pairs of



strange particles are of an entirely different nature. There are no particle-antiparticle groups here. For instance, let us take the reaction

$$\pi^- + p \rightarrow \Sigma^- + K^+.$$

A negative pi-meson collides with a proton generating a negative Σ-hyperon and a positive K-meson. The Σ^- particle further decays into a pi-meson and a neutron that does not leave a trace in the chamber. (In our case, the pion then entered a carbon nucleus and destroyed it.) The Σ^- and K^+ particles are in no way connected with the particle-antiparticle relationship. The situation is much the same also in the other reactions of strange-particle production.

Why? The answer lay outside the limits of existing theory.

However, the whims and freaks of production processes were not all. The new particles decayed in a still stranger fashion.

Take another look at the above reaction. It involves a proton and a pi-meson, which are obviously strong-interacting particles. Hence, also the other two particles—both Σ-hyperon and K-meson—should be considered as belonging in the strong-interaction category.

This is indeed corroborated by a whole series of both theoretical and experimental material: for instance,

hyperons could readily take the place of nucleons in the nucleus if they were not so unstable (which was mentioned when we were dealing with hypernuclei).

Thus, hyperons (for the sake of brevity we shall confine ourselves to them) are strong-interacting particles. This is also quite in accord with the fact that they are produced very "actively". That being so, they should readily eject pi-mesons and convert into nucleons, say according to the following reaction:

$$\Lambda^0 \rightarrow p + \pi^-,$$

which, by the way, is actually observed.

However, the amazing thing is that this process is radically inhibited for some reason. And other processes are too. Insofar as hyperons are strong-interacting particles, they should all decay almost as soon as they are born. It should take just about the amount of time required for a ray of light to cover a distance equal to the size of one particle (light, you recall, takes only a tenth of a second to circle the equator of the earth). What does experiment say? Experiment says that hyperons live hundreds of millions of millions of times longer than befits strongly interacting particles. Strange? Certainly, but then the particles are called "strange particles".

Strangeness has not yet been exhausted. Compute the charge responsible for hyperon decay and you will get something fantastic (though by this time we are all quite used to such things): in place of the strong-interaction constant we get—with convincing precision—what do you think? *The constant of the weak interaction!*

As Gell-Mann and Rosenbaum say, the strange particles move away from each other after birth and away from death by means of the strong interaction, and they continue to exist until a less probable weak process finishes them off.

The longevity secret of strange particles

There is no doubt, therefore, that some kind of factors prevent the "strong" decay of strange particles. Long experience has taught physicists to seek laws of conser-

vation when a restriction of any kind appears. Transformations are forbidden if the law of conservation of charge is violated. The law of conservation of energy prohibits processes in which the overall mass of decay products is greater than the mass of the decaying particle. The laws of conservation of energy and momentum insist that when an electron-positron pair is annihilated no less than two gamma quanta are born.

Perhaps this retarded strong-decay of hyperons signifies a new and as yet undiscovered conservation law.

That was the hypothesis advanced by Gell-Mann. The new quantity conserved in strong and electromagnetic interactions was called "strangeness".

To the ordinary particles—the proton, neutron (and their antiparticles)—and also to the neutral and charged pi-mesons we attribute zero strangeness. For the other strong-interacting particles the strangeness is distributed as follows:

strangeness equal to minus unity—Λ^0, Σ^+, Σ^-, Σ^0, K^-, $\widetilde{K}^0$,
strangeness equal to plus unity—$\widetilde{\Lambda}^0$, $\widetilde{\Sigma}^+$, $\widetilde{\Sigma}^-$, $\widetilde{\Sigma}^0$, K^+, K^0 (the corresponding antiparticles),
strangeness equal to minus two—Ξ^-, Ξ^0,
strangeness equal to plus two—$\widetilde{\Xi}^-$, $\widetilde{\Xi}^0$.

If the reaction is such that the strangeness does not change, then the process is a strong one. For instance, in the above case of

$$p+\pi^- \rightarrow \Sigma^- + K^+$$

the initial particles have zero strangeness; Σ^- has strangeness -1, K^+ has strangeness $+1$. Consequently, the total strangeness on the right side is also zero. The "strong" pattern is solved.

On the other hand, the decay

$$\Lambda^0 \rightarrow p+\pi^-$$

occurs with an obvious change in strangeness of unity (from -1 to 0). According to the new conservation law, this transformation cannot occur by the laws of strong interaction. Pair production is determined by the conservation of strangeness. Like the Soviet law for mountain climbing prohibits solo ascents (there must be at least

two persons), so the law of conservation of strangeness produces hyperons and K-mesons only in pairs. If it weren't for this law, particles would perish immediately after their birth.

To get a better picture of what this is all about, imagine that a certain process of the strong category (that is, without any change in strangeness) lasts one second. Then transformations involving change in strangeness by unity would require millions of years! For a change of two units of strangeness, incredibly large periods of time—far greater than the age of our earth—would be required.

Today, we are very far away from the charming simplicity of the physical picture of the world visualized by scientists at the start of the century, when the whole world was made out of two kinds of particles—electrons and protons (with photons added). We now know that the world is much more complicated than the picture of only twenty years ago. And this concerns not only the world of elementary particles in the direct sense of the word. Changes have taken place in our conceptions of the past and the present of the macroworld, the world of cosmic objects.

3

The neutrino in the universe

Take the role of the neutrino in the universe. We are already used to this particle invading every imaginable sphere. A number of stellar processes are accompanied by the emission of neutrinos. The reason is obvious since the main source of stellar energy is from nuclear fusion. The chain of processes resulting in the conversion of four protons into an alpha particle (a nucleus of helium) gives rise to two positrons; and this is always accompanied by the ejection of two neutrinos. Positrons and electrons are

annihilated and a neutrino escapes from the star. From the sun alone, over 10^{11} neutrinos strike every square centimetre of the earth's surface every second.

Add to this the fact that in interstellar space mesons and hyperons are undergoing decay. And all these disintegrations are accompanied by the generation of neutrinos and antineutrinos.

Deeper in stellar interiors, under great pressures and at very high temperatures we find the direct production (by electrons) of neutrino-antineutrino pairs, probably in the presence of nuclei:

$$e^- \rightarrow e^- + \nu + \bar{\nu} \quad \text{or}$$

$$e^- + \gamma \rightarrow e^- + \gamma + \nu + \bar{\nu}, \text{ etc.}$$

Due to such processes, the neutrino luminosity of stars should exceed their optical luminosity.

What happens to the neutrinos then? These all-penetrating particles naturally pass right through enormous stellar thicknesses and carry away their portion of energy even if they got started deep inside the star. Given the low density of the universe, neutrinos should be able to fly for eons without ever being absorbed. That is why there should be an *accumulation of neutrinos in the universe.*

What is more, it might be (as we shall see later on) that a considerable portion of the universe was made up of neutrinos in the early stages of expansion. In the main, these neutrinos are still intact, having lost a considerable portion of their energy during the expansion of the universe.

This reasoning leads us to presuppose enormous quantities of neutrinos in the world: the neutrino mass might even compete with the entire mass of visible matter. If that is so, the density of matter of this world, and hence its curvature, is largely determined by neutrinos. It may be that neutrinos make our universe finite.

If only we could construct neutrino-receiving sets of the same sensitivity as radio-receiving sets. What fantastic facts we would be able to extract. We might know, today, whether our universe is finite or not. We would be able to glimpse into the very centre of the largest stars. Remember, neutrinos are born in those parts and carry away some sort of "reminiscences" about the processes

that gave them birth. The neutrino would permit penetrating whole galaxies much better than X-rays pierce a sheet of paper.

Incidentally, even without neutrino-receiving sets we have been able to learn a great deal about the evolution of the universe and the stars, thanks to these particles.

Stellar evolution

Nuclear reactions are what make stars burn continuously. They feed on hydrogen, converting it into helium, and this is accompanied by a release of stupendous energies. But the fuel supplies of any star are limited. What happens after the hydrogen has burnt out?

In the stationary state, when the fuel supplies are still great, stars are in equilibrium because the gravitational pressure that tends to compress the star is balanced by the pressure of gas particles that constitute it (at high stellar temperatures, matter is in an ionized-gas state). This pressure is proportional to the temperature. As the hydrogen burns away and the temperature falls in the central regions of the star, the pressure will diminish and the star will begin to shrink. An interesting thing is that in the process the peripheral areas will heat up due to the fact that the surface of the star diminishes and radiation from the interior coming to unit surface increases.

As a result, if the mass of the star is less than 1.4 solar masses we have a stable configuration known as the white dwarf. The best known white dwarf is the satellite of Sirius, a tiny star lost in the bright rays of its primary.

White dwarfs are a possible final stage in stellar evolution. The luminosity of such stars is very low, but with a mass of the order of our sun their radius is only roughly that of the Earth or Uranus. Accordingly, white dwarfs have stupendous densities: 10^8 g/cm^3. Under these conditions, the atoms are totally ionized and the star consists of tightly packed nuclei and electrons.

As the Soviet physicist Landau has shown, white dwarfs do not represent the only possible configuration of a star after the nuclear fuel has all burnt out. Under the great compression of a star, the electrons can be "squeezed" into their nuclei and, via the reaction $e^- + p \rightarrow n + \nu$, convert protons into neutrons. Neutrinos leave the star,

and the neutrons remain behind. Now the decay of the neutrons according to the ordinary scheme

$$n \longrightarrow p + e^- + \bar{\nu}$$

is impossible because of Pauli's principle. Here is how matters stand.

The domain of electron motion is restricted to the dimensions of the star. Quantum mechanics states that when motion is restricted, electron energy cannot be arbitrary. As in the atom, only distinct discrete values of energy are possible. If we picture a star in the form of a box filled with electrons and other particles in such a way that the more energy a particle has, the farther it is away from the bottom of the box, we should imagine this box to have a multitude of separate shelves, each one corresponding to a definite particle energy. There are shelves for electrons, protons, and so forth.

According to the Pauli principle, two electrons cannot be in the same state at the same time. For that reason, one electron shelf accommodates only two electrons with opposite spins. At not too high temperatures all the lower shelves (up to a certain one corresponding to the maximum energy E_m of the electron) will be filled. Professionally speaking, the electron gas will be in a state of degeneration with the lowest energy allowed by quantum mechanics.

Now a neutron cannot decay if the energy resulting from electron disintegration is less than maximum E_m. This is because all the lower energy shelves are occupied and there is no place for fresh electrons in the star. It's almost as if a neutron "knew" that it couldn't decay without violating the laws of quantum mechanics, and so remains stable.

The result is that an ordinary star evolves into a stable neutron star with a slight admixture of electrons and protons. Densities here are fantastic: 10^{14} g/cm^3. This is nuclear density. *

As in the case of white dwarfs, stellar compression should at first be accompanied by an increase in surface temperature. But now, for some time, the temperature will rise to such an extent that the star will begin to radiate X-rays and gamma rays. It is interesting to note that

* As the Soviet astrophysicist Ambartsumyan has demonstrated, hyperon stars are also possible.

the first (though not fully reliable) indications of the existence of cosmic sources with this type of radiation were obtained with the aid of artificial satellites.

A neutron star will have a stable configuration if its mass does not exceed twice the solar mass. Now what will happen if the mass is greater?

According to a theory proposed by Oppenheimer and Sneider, the star will begin to shrink without letup, passing beyond its gravitational radius. This is called "gravitational collapse". The gravitational radius determines the critical dimensions of a body and depends on the mass m and also on the velocity of light c and the gravitational constant $\varkappa$: $r_0 = \frac{2m\varkappa}{c^2}$. As follows from Einstein's theory of gravitation, all radiation and signals inside a sphere bounded by the gravitational radius will be propagated only towards the centre.

To an outsider, this process will appear as follows. At first the star contracts with extreme rapidity and then more slowly. The luminosity of the star falls off rapidly and then ceases altogether. Light is no longer able to overcome the gravitational pull and escape the star. A star of mass greater than twice that of the sun turns into a black sphere just a few kilometres across. It is identifiable only by its static gravitational field.

We still do not know whether such stars exist in the universe, and if so how many there are.

An early stage in the evolution of the universe

When discussing the expanding universe we touched on one of the most cardinal of problems: the constituents of the universe at the time of maximum compression. The initial state of the universe must be such as to account for the present-day universe.

It is most natural to assume that in the ultra-dense state, the universe consisted of cold neutrons, which would not decay for the very same reason that they are stable inside a neutron star.

During the initial moments, the universe expanded with extreme rapidity; 15 minutes after commencement of

expansion, its density should have equalled that of water. During expansion, the energy shelves for electrons were arranged in accord with the laws of quantum mechanics closer and closer together, so that very soon the electron energy at the highest occupied shelf was less than the energy of an electron produced in neutron decay. For this reason, neutrons would begin to disintegrate and protons would appear.

Proton-neutron collisions give rise to nuclei of heavy hydrogen: deuterons. Deuterons collide with one another and with protons to form helium nuclei and tritium (extra-heavy hydrogen). As a result, the universe would very soon be poor in protons. Actually, however, during the first stages of evolution, matter in the universe consisted mostly of hydrogen (90%). And the stars shine brightly at present precisely because of large supplies of hydrogen. That dispenses with the neutron hypothesis of the primordial state of the universe.

At present the most promising theory appears to be that of Zeldovich. According to this theory, at the start the universe was very dense, cold and consisted of protons, electrons and neutrinos. All possible lower energy levels of neutrinos were completely filled. This was because neutrinos could not leave the universe the way they escape from a star.

For that reason, the Pauli principle prohibited the transformation of protons into neutrons via the reaction.

$$p + e^- \longrightarrow n + \nu.$$

There simply was no room in the universe for another neutrino, just as there is no place for a third electron in the first shell of an atom. The presence of neutrinos stabilized protons at high densities.

As the universe expanded, there was a fall in the probability of electron-proton collisions and, hence, the formation of neutrons in considerable quantities.

Towards the end of the initial, proto-stellar, state of the universe, ultra-dense matter converted into pure hydrogen almost completely. It was only later, in stellar interiors, that the chemical elements began to emerge.

According to the Zeldovich hypothesis, neutrinos play one of the most significant roles during the first act of evolution of the universe.

Just a little over three decades have passed since the appearance in science of the term "neutrino" that sparked a marvellous series of discoveries which we associate with the weak interaction. The study of these interactions played another and no less important part. The ocean of the unknown that Newton spoke of three hundred years ago is more keenly felt today than ever before.

An Early Summary of What We Have Learned

The magnificantly integrated picture of interactions that confront the investigator today—four basic types—displays an infinitely diverse range of processes. Though extremely different, these four basic types of interaction have a number of fundamentally unifying characteristics.

Unlike? Certainly, for there is a great difference between gravitational and, say, nuclear forces. Their ranges of action have nothing in common. The "spheres of influence" of each force appear at first glance to be well defined. Gravitational forces predominate in the world of cosmic entities. The sphere of electromagnetic forces involves atoms, molecules and pieces of matter made up of them. Geometrically, the scale is not cosmic, but it is fantastically rich and diverse, ranging from the spectra of the simplest atoms to the extremely complex processes that occur in living organisms. The realm where nuclear forces dominate is smaller still: atomic nuclei. And, finally, the weak interaction, which determines the processes occurring in the intimate sphere of individual particles, including the atomic nuclei that make up matter.

This gives us our first rough classification of forces and their range of action:

COSMOS—ATOM—NUCLEUS—PARTICLE.

The second feature is the magnitude of the force or, more precisely, the magnitude of the energy that corresponds to different interactions. The nuclear interaction exceeds the electromagnetic interaction by about one hundred times; and it is 100,000,000,000,000 times the weak interaction. The gravitational interaction of two electrons

is so small compared with the Coulomb interaction that we would have to write a number followed by 44 zeros.

Much more could be said about the dissimilarities of the fundamental forces. Yet no matter how different they appear, there is a very profound unity in all these forces.

For instance, the division of spheres of influence. It is not so absolute after all, for when dealing with cosmic bodies, one cannot entirely dispense with nuclear forces, or electromagnetic forces, or, finally, the weak interaction. To do that would mean crossing out all of stellar physics.

Or take this. Nuclear forces are not the only essential ones in a nucleus. Electromagnetic and weak interactions are important here as well. Even gravitational forces with their classical sphere of action in the cosmos can, according to certain investigators, play important roles in the microworld in the formation of the particles themselves.

Thus, pigeonholing the different types of interaction is not enough, for the physiognomy of the world is determined by the totality of their actions, their integration and profoundly harmonious interrelations.

The world is a unified entity, and this unification refers also to the world of interactions.

And one thing more. Why do we speak only of four types of forces? Is there any assurance that new types of interaction will not be discovered in ten or twenty years?

Naturally, there is not. We have described the forces which today appear to be fundamental. But even in this book we have insinuated that the family of forces might be expanded a bit. Recall, in the chapter on nuclear forces, the problem of repulsion of protons and neutrons at ultra-small distances. Does this repulsion involve only one of the four fundamental forces? As far as we can see, the answer may well be No. Perhaps a new member will appear in our family right at this spot.

And this isn't all. Suggestions have been made that there might be forces of other types, say, specific meson interactions and other same-particle interactions that are often called direct (that is, without mediating particles). It may be that these forces will soon lengthen our present list of forces.

On the other hand, something quite different might occur. In the search for deeply hidden relationships between natural phenomena, we might come upon the fea-

tures of a new unity that evades us today. And who knows but that this unity might not help to interpret nuclear and weak interactions (and still other as yet unknown forces) as a manifestation of certain general laws governing matter.

However, there are still more attractive possibilities. To explain a process or physical phenomenon, we must in principle know the structure of matter, the equations of motion and the forces. Specific features of a very profound unity are emerging within these basic elements of the physical theory.

Unity in the structure of matter consists, above all, in the fact that the entire world is built up of a comparatively small number of elementary particles. Enough has been said of the unity of forces, which is really the substance of this whole book. Finally, there is a definite unity in the equations of motion. Undoubtedly the highest form of unity that we can imagine is a merging of all three basic elements of the physical theory. A unification like this is possible, as witness the general theory of relativity of Einstein. In this theory, the equations for the gravitational field also determine the trajectory of motion (in this field) of any point mass.

In the unified quantum field theory now under construction there is a new kind of underlying equation of motion. This equation is written for a unified field (matter), all possible forms of existence of which are the various elementary particles. It is hoped that this equation will determine what elementary particles can exist in nature and will give a mass spectrum of the particles. This equation should be able to describe the interacting particles straight off. The theory does not countenance particles that do not interact with anything. In fact, no such particles are known. This extensive program, which is being developed by Heisenberg and other scientists, at one time appeared to offer considerable promise. So far however no significant advance has been made.

The program of a unified field theory is attractive in many ways, but there can be no guarantee of success. Science never offers guarantees, anyway.

If a unified field theory is ever constructed, a new book on the forces of nature will be quite different from this one. But different in what way we can only guess.

C H A P T E R S I X

Nothing gives such weight and dignity to a book as an Appendix

(Herodotus)

IN LIEU OF AN APPENDIX

1 *What Are Resonance Particles*

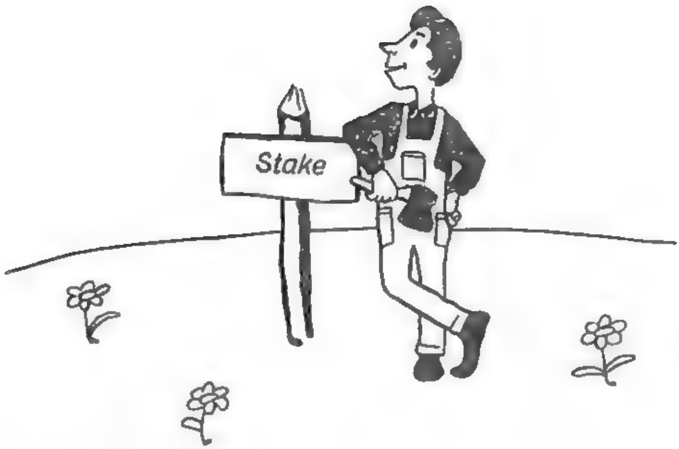
2 *Systematics of the Elementary Particles*

1

All about why a new chapter appeared in this book of the forces of nature

The Soviet scientific journal *Advances in the Physical Sciences* of March 1965 carried an article entitled "Principal Trends in Elementary Particle Physics" by A. Baldin and A. Komar, who summarized the results of a high-energy physics conference at Dubna. Speaking of types of interaction, they posed the following question: "Is their multiplicity exhausted by strong, electromagnetic, weak and gravitational interactions, or is such a division arbitrary and does there exist a multitude of intermediate types of interaction, or, finally, are they all diverse aspects of a single interaction that embraces all the elementary particles?"

The very fact that such a question is posed may surprise or even vex the reader, for everything so far has indicated a clear-cut system of four types of fundamental forces.



We have said that in the vast expanses of the universe, on our own planet, in any piece of matter, in living organisms, in atoms and their nuclei, and, finally, in the mutual transformations of the elementary particles we encounter only four types of force. That was the situation just recently, while chapters one to five were being written. Now, since the manuscript went to press, things have changed somewhat.

Fresh ideas are coming into physics and a new chapter is needed on the unity and diversity of nature. The mysterious and tightly closed world of elementary particles has opened up, though not everything has been tested. Yet physicists have staked a new claim in a vast unexplored territory. The claim is serious, too, for we already have the prediction and discovery of a new particle. This is important because we are dealing with a system for ordering particles, the systematics of elementary particles.

As always, every new discovery not only resolves a problem or two, but gives rise to a multitude of fresh problems. That is exactly what happened this time. The new discoveries brought forth the questions that we mentioned in connection with the paper of Baldin and Komar. These and many others. They are problems which physicists often term “Problem Number One”.

We have decided not to change anything in the preceding chapters and to add this last discussion of recent advances. This may produce a certain disharmony in the

book, but then the reader will better feel that science is not a rigid system of dogmas but rather a living, developing and often intrinsically contradictory organism. Which is all to the good.

It could easily be that the material of this appendix will tomorrow become a chapter in a new book: new yet old—a great book of the cognition of nature.*

Elementary particles...

What are they?

The term "atom" has been bandied about in thousands of books since the days of the ancient Greek scholar Democritus. It means indivisible. Atoms are rejected, doubts are raised, then atoms are believed in, but, curious as it may seem, atoms do not actually justify their name. The atoms that every schoolchild knows about have a well-studied structure that can be analyzed into component parts.

Atoms are divisible and their division (splitting) has become a decisive factor in our life. Atoms consist of smaller particles. Perhaps we should extend the term "atom" to these particles. Actually this was done, only the terminology was altered somewhat: the particles became known as elementary particles.

Now just what is an elementary particle? First of all, elementary is rather ambiguous as it is, meaning both something that is grasped at once and something so fundamental as to be beyond the understanding of anyone. It is in this latter sense that the subatomic particles are called elementary.

To begin with, everything appeared quite simple: the elementary particles were the ultimately simple and indivisible units of matter. The discovery of every new particle was (and still is) a remarkable triumph of science. But for the past twenty years or so with every fresh triumph there has been a good deal of anxiety: the triumphs have been just a little too frequent. The number of particles had already exceeded thirty. Were they all elementary?

* N. P. Yudin is co-author of Chapter Six.

They included some, like the muon, that are still complete freaks of nature. The world could easily get along without them, one would think.

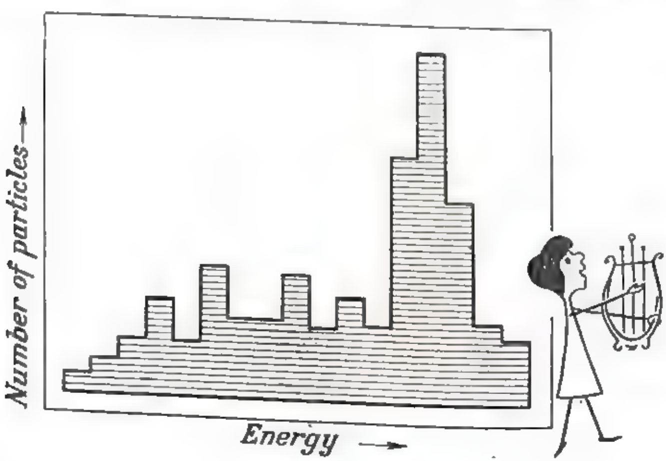
Then another thing. Most particles are not eternal entities. They appear, live a short time measured in minutes for the neutron and small fractions of a second for the pi-neutral meson and other particles, and then vanish giving rise to new particles.

Still, after much hesitation, scientists decided that all particles are elementary and their demise is to be considered not a decay into constituent parts but a transformation into other particles. The point is that particle offspring differ from their particle progenitors in quite another way than the fragments of a broken jar from the original jar.

The shortest-lived of all

The situation in the physics of elementary particles was extremely tense when, 2 or 3 years ago, an event occurred that aggravated it still more. A new series of particles was discovered with such short lifetimes that by comparison one of the short-living particles like the pi-neutral-meson would appear immortal, for its longevity is a thousand million times greater.

The lifetime of these particles (10^{-23} sec) is such that they do not leave any traces in a cloud chamber. They come to life and then die so quickly that they hardly



have time to traverse a distance the size of a proton (10^{-13} cm).

The question immediately arises as to whether we should include such particles in the list of elementary particles or not. On the one hand, we ought to, since, like ordinary particles, they have mass, charge, a lifetime, spin, and so forth. But they have such short lifetimes. The amazing thing is that scientists were able to detect them at all. Still and all, the discovery was made and done very convincingly. Elusive though they are, they were caught.

The name of such particles—resonances or resonons—speaks more of the methods by which they were discovered than of their nature.

Catching the elusive particle

The first group of resonance particles was discovered in the study of the scattering of pi-mesons on nucleons (protons and neutrons). Scientists bombarded a hydrogen-containing target with a beam of positive pions and found particularly intense scattering at a kinetic pion energy of 200 million electron-volts. In physical parlance, this is called a resonance: the number of scattered mesons increases sharply at a definite energy.

The word "resonance" is probably familiar. A tuning fork is in resonance (i.e., oscillates intensely) when a sound wave falls on it whose frequency of vibrations coincides with the natural frequency of the tuning fork. This is typical of wave scattering of any kind.

Now recall that in quantum mechanics a very simple and general relationship is established between frequency and energy. The energy differs from the frequency only by the factor h (Planck's constant). Which means that in quantum language, resonance fits the case when the energy of the scattering particles (they are also waves according to wave-particle duality) coincides with the energies allowed for the scattering agent.

We can continue our analogy. When a sound wave hits the tuning fork, the prongs are at first stationary and then begin to vibrate. If the frequency of the sound is far removed from the natural frequency of the tuning fork,

the vibrations will be weak and will die out if not reinforced from without. The situation is quite different in the case of resonance. Here the vibrations are much more stable and continue for quite some time even if there is no reinforcement. We might say that resonance occurs when a state arises in scattering that is itself relatively stable.

We are not quite sure as yet just how scattering takes place (of pions on protons, say). But the very fact of resonance inevitably indicates that during the time interval in which scattering occurs some kind of stable formation (relatively stable, naturally) manifests itself. Ordinarily, scattering is symbolized as follows:

$$\pi^+ + p \rightarrow N^{*++} \rightarrow \pi^+ + p.$$

It reads: a positive pion collides with a proton to form an intermediate state N^{*++}, which disintegrates into a positive pion and a proton. The most important quantity characterizing the intermediate system N^{*++} is its energy (or, by virtue of the familiar relationship $E = mc^2$, the mass).

Thus, scattering exhibits some sort of relatively stable formations that have definite parameters (mass and, as is readily seen, charge, etc.) and are so much like ordinary particles in every way that we have no grounds to reject them as particles. Now what is this intermediate state N^{*++}?

One can suppose that a pion and proton are capable of forming a system reminiscent of the ordinary hydrogen atom without losing their individuality. This is a "pi-mesonic atom" in which a meson takes the place of the electron, and the role of Coulomb forces is played by nuclear forces. As in the conventional atom, the energy of the pi-mesonic atom must be quantized, that is, it has to take on a discrete series of energy values.

The probability of the formation of an "atom" and, consequently, the probability of considerable scattering of a pi-meson after decay of the "atom" will be maximal for the case when the energy of the incident meson is exactly equal to the energy of the "atom". That is when resonance is observed.

We can estimate the lifetime of the "atom" from the width of the resonance curve. Of help here is the Heisen-

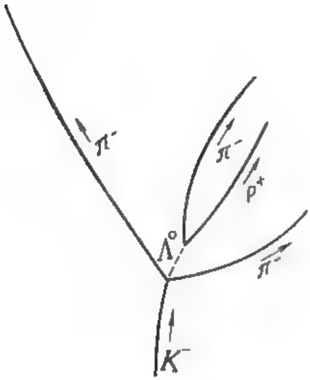
berg uncertainty relation between energy and time. The width of the curve yields the order of the uncertainty of energy of the "atom", ΔE. Its lifetime is $\Delta t = \frac{h}{\Delta E}$. It comes out to roughly 10^{-23} sec. Moving with a velocity close to that of light, the meson covers a distance of about 10^{-13} cm, which is exactly equal to the dimensions of the region of meson-proton interaction. Thus, the lifetime of the "atom" is such that the pi-meson hardly has time to execute a single revolution about the proton. Such being the conditions, it is rather far-fetched to call this system an atom.

On the other hand, we can presume that the pi-meson and proton merge for a certain time to form a new particle, which then disintegrates once again into a pi-meson and a proton. It is easy to determine the mass of this particle from the laws of conservation of energy and momentum. It comes out to 1,237 million electron-volts if we express the mass in units of energy, as is frequently done now in the physics of elementary particles.

Without attempting to determine the true nature of this intermediate state N^{*++}, physicists have cautiously termed it a resonance particle, in this way stressing the rather obscure nature of this formation.

Where can we find a hyperon target?

When resonances associated with the scattering of pi-mesons on nucleons were discovered, hardly anyone doubted that resonances are not peculiar to these systems alone. And true enough, resonances were also found in meson-hyperon systems (baryon resonances). A group of meson resonances has also been discovered ($\pi-\pi$, $\pi-K$, $K-K$, etc.). But in these cases they cannot be detected through studies of particle scattering because one cannot make a target out of, say, Λ^0-hyperons. Hyperons decay during a time on the order of 10^{-10} sec. And pi-mesons disintegrate as well. How resonances are detected in such systems will be shown for the case of resonance in a system of a pi-meson and a Λ^0-hyperon.



If protons are bombarded with K-mesons of high energy, we often get the following reaction:

$$K^- + p \rightarrow \Lambda^0 + \pi^+ + \pi^-.$$

A K-meson collides with a proton, produces a Λ^0-hyperon and two pi-mesons.

By studying a sufficiently large number of such reactions, one can find the number of negative pions having a definite energy. We can then construct a curve of the number of negative pions as a function of their energy: a so-called meson-energy spectrum. The character of the curve should depend on how the given reaction develops. Let us assume that in the reaction all three particles (Λ^0, π^+ and π^-) are born at once and fly out in different directions independently of one another. Then the initial energy of the K^--meson and the proton will be redistributed in different ways between the generated particles. The laws of conservation of energy and momentum uniquely determine only the maximum possible value of energy of the negative pion. Energy can take on any values from zero to maximum. The situation will be different if the Λ^0-hyperon and positive pion behave as an integral entity immediately after the reaction. Then the initial energy of the K^-- and p-particles will be distributed between two particles and the laws of conservation of energy and momentum will uniquely determine the values of energy of both emerging particles. To put it rather crudely, the laws of conservation of energy and momentum are in this case two equations in two unknowns (the energies of the new particles) since the momentum may be ex-

pressed in terms of the energy. But if three particles are born at once, there would be three unknowns and the energies of the newly born particles could not be determined uniquely.

The experimental curve has a sharp pip (peak) for a definite energy value of the negative pion. Which means that in a large number of cases not all three particles are born at once. First two are born and then one of them disintegrates:

$$K^- + p \rightarrow Y^{*+} + \pi^- \rightarrow \Lambda^0 + \pi^+ + \pi^-.$$

Y^{*+} is an intermediate system that behaves like an integral entity. Like N^{*++}, it is called a resonance particle. Using the conservation laws, we find its mass: 1,389 million electron-volts. From the width of the resonance peak we get the lifetime of the resonance Y^{*+}. It is exactly the same as that of the N^{*++}.

Besides the positively charged resonance Y^{*+} we have a negatively charged Y^{*-} and a neutral Y^{*0}. Their masses are about the same.

The other resonances were discovered similarly. There are now about fifty different resonances.

At the International Conference on High-energy Physics that took place in August 1964 (Dubna), the participants were given tables of baryon and meson resonances with the latest and most complete data. The first one is given at the end of this book to illustrate the present situation.

2

One elementary particle as a composite of all the others

The discovery of resonances greatly complicated the already complex picture of the world. The question of what an elementary particle really was became acute.

One could not be sure that all 32 particles with their relatively long lifetimes (some of them infinitely great) are elementary. There were simply too many of them. And then all of these resonance particles.

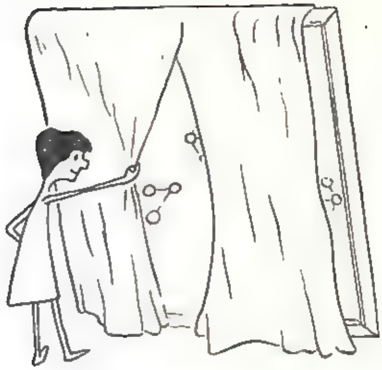
But that is not all. It is now quite clear that the particles which we call elementary actually possess an extremely intricate structure. For example, protons and neutrons are surrounded by clouds of pions. Pions are also structurally a part of the nucleon. In turn, pi-mesons can convert into a nucleon-antinucleon pair. Hence, particles cannot be pigeonholed and studied separately. To describe one particle, we have to take into consideration the others too. As Markov has put it, each elementary particle begins to be pictured as a composite of all the others. So the circle closes: the properties of one of the particles are determined by all the others. Does the circle open at any point? In other words, are there ultimately elemental particles that go to build all the others, or does nature lack such entities? We do not know today. There is no criterion of elementariness of a particle.

We may hope that the discovery of the resonances has only temporarily complicated the picture of the world. Subsequent studies may help to solve the problem of the internal structure of the elementary particles—Problem No. 1 of modern physics. It is precisely in the problem of resonance that we see the most obvious necessity to change the methods of theoretical description that have been used so far.

Faint signs of a unity

The discovery of resonances has already helped us to see certain things that have been hidden. Imagine that you are confronted with a large canvas covered with a heavy curtain. You want to get a peek at the picture but you are unable to pull the curtain aside . So you tear a piece in the middle and note some figures. Now you are positive that there is a real picture with many participants behind the curtain and not simply a blank canvas.

Another jerk and you see whole portions which indicate that there is a profound affinity between the various



parts, an underlying unity of composition. The other portions are so small that it is hard to understand how their various parts are connected.

The scattered fragments do not as yet indicate what unity there is to the picture, but gradually, as the curtain falls, we recognize the remarkably integral character of the picture. The separate pieces fall into place and a full picture begins to emerge.

The picture is not simple; in fact, it is a whole series of pictures more involved than the classical triptych.

Something like this occurred just recently when the remarkable world of the elementary particles was reviewed after the discovery of resonances.

Twins and triplets in electric suits

When physicists just began to learn about the elementary particles, each one of them appeared as a brilliant and isolated individuality. But they grew in number and soon, of themselves, began to separate into groups having similar features.

Classification of the particles began. But what features were to be used in such systematics? Historically the first was the mass. (Perhaps Mendeleyev's classification of the elements played no small part in this respect.) The particles were divided into groups, as they were called: light, medium and heavy. However, this scheme re-

quired certain alterations. Spin had to be taken into consideration and—so important to us—the type of interaction as well. That produced the familiar classes of particles known as leptons, mesons *, and the heavy particles, or baryons, and, finally, the resonances were added. Only the photon remains separate forming a whole class by itself **. We examined this classification when analyzing the table of elementary particles.

There is a curious thing here too. A close glance at the table will show us that there are certain groups of particles among the mesons and baryons that conveniently form subgroups.

Take the three pi-mesons: π^+, π^- and π^0. If it weren't for the charges, how would we distinguish them? They have absolutely identical behaviour as far as strong interactions go. The spin is the same. Even the slight difference in mass is of electromagnetic origin; their masses would be the same if the electric charges disappeared. One feels inclined to say that the three pi-mesons are actually one particle, not three, and simply have different states of charge.

The pions are not the only set of particles that have only an electrical difference. The reader will remember that our discussion of nuclear forces suggested the same thing for the proton and neutron. We might say that this is the rule and not an exception. Which is food for thought.

Let us take a closer look at charge multiplets (the name given to groups of particles that differ only as to states of charge). In addition to the triplet (a multiplet of three particles) of pi-mesons, there is a triplet of Σ-hyperons. K-mesons, like the proton and neutron, form a charge doublet (two particles); the Λ^0-hyperon forms a singlet, which means the sole representative of a special charge multiplet.

Twins and triplets in electric suits are also found among the resonances. For instance, in addition to the resonance N^{*++} that appears in the scattering of positive pions on

---

* The big mass of the mu-meson deceived scientists into placing it among the mesons. It is now clear, however, that it is by nature a lepton.

** Attempts have been made to establish an affinity with the neutrino, but without success.

protons, we also have the pion-nucleon resonances N^{*+}, N^{*0}, and N^{*-} that differ solely in electric charge. The charges are $2e$, e, 0, and $-e$, respectively.

The most important stamp in the passport of an elementary particle is its membership in a specific charge multiplet and the number of particles in that multiplet. However, it appears much more convenient to speak not of the number of particles in the multiplet, but of the so-called isotopic spin and of the projections of isotopic spin.

What is isotopic spin?

The term isotopic spin is a combination of concepts that we have already encountered separately. Isotopes, as you recall, are twin elements that do not differ chemically but have somewhat unlike physical parameters, such as mass, for instance.

In the Periodic Table, all isotopes of an element reside in a single square and all have the same name. We say helium-three (He^3) and helium-four (He^4) are not different elements but, if you like, different states of one and the same element. Likewise, all particles that are members of a single charge multiplet family are considered to be one particle in different states.

Then where does spin come in? If we mean the spin we discussed in Chapter 3 and described as intrinsic angular momentum, it doesn't come in at all. If one term is used in quite different meanings, there is nothing new or surprising either. The word degree is used to designate temperature and to measure angles. And angles are measured in minutes and seconds just like time. Incidentally, this latter case deserves more attention. The second hand of a watch glides round the dial marking off one hour in one circuit; now a change of angle in one minute of time is exactly equal to one "angular" minute. The parallel is complete, and is of course associated with the mechanical model that we use to measure time (an hourglass would yield nothing in that respect).

Now let us take another "mechanical model", a particle with a definite spin. Let this spin be, say, 1/2 (all

in units of $\hbar$, remember). As you recall, a particle of this kind can have only two orientations: its spin is either parallel to the momentum, or antiparallel. Two orientations, two possible states. And if the spin is zero, then naturally nothing changes and there can only be one state.

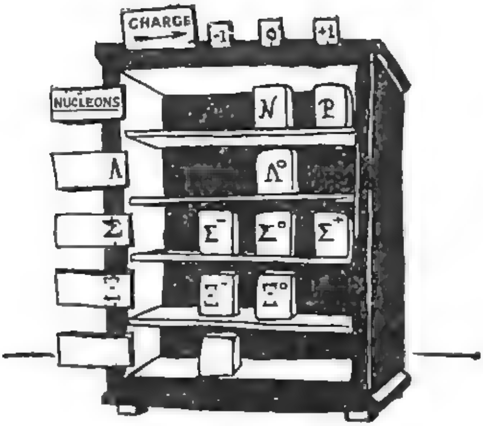
But what about spin one? A quantum calculation indicates that three different states are possible. If the particle spin is 3/2, then there are four such states, and so on. In the general case, spin equal to n yields $2n+1$ different "internal" (as they are called) states. Now let us go back to our "families" of particles. It will be recalled that we regard the members of each such family as a single particle in different internal states. Now the number of such states varies with the family. The poorest (Λ^0-hyperon) has only one (Reminiscent of zero spin, isn't it?). Others (nucleons) have two. Which is again a complete analogy with the case of spin 1/2. Pi-mesons (π^+, π^0 and π^-) have three, as in the case of spin unity, and so forth. Now you see that our mechanical model has fully justified itself, so there are grounds for speaking of isotopic spin. But we can take another step. Recall that whereas the value of spin (we are again speaking of mechanical spin) shows the number of possible internal states, the concrete specification of each of these states may be attained, for example, by indicating the sense of the spin (with respect to the momentum or in general to any axis; traditionally, the z-axis is chosen). The same can be expressed in the following form as well: the state is determined by the projection of spin on the z-axis. Nothing prevents us from doing likewise in the case of isotopic spin as well: we assign to each of the members of our "iso-spin families" a definite value of the "projection of isotopic spin on the z-axis". One, of course, might simply indicate the electric charge, because from the foregoing it is evident that this projection is what determines the charge of the particle of the family; however, it turns out that precisely the projection of isotopic spin is the only acceptable quantity for the mathematical apparatus of the theory.

In the table we find several instances of correspondence between the number of particles in a charge multiplet and the charges of the particles, on the one hand, and the isotopic spin together with its projections, on the other.

But a picture is better than a story. Imagine that we are engaged in particle systematics (producing baryons, for instance), i.e., we put them in pigeonholes in a very literal sense. Let us draw the holes—one for each charge multiplet—and allot places for the particles in strict accord with their charges.

Now that everything is arranged we can see a certain asymmetry in the arrangement which may be described quantitatively. All we have to do is determine the mean

Particle	Number of particles in multiplet	Isotopic spin	Charge of particle	Projection of isotopic spin
Λ^0	0	0	0	0
p	2	$\frac{1}{2}$	$+e$	$+\frac{1}{2}$
n			0	$-\frac{1}{2}$
π^+	3	1	$+e$	$+1$
π^0			0	0
π^-			$-e$	-1
K^+	2	$\frac{1}{2}$	$+e$	$+\frac{1}{2}$
K^0			0	$-\frac{1}{2}$
Σ^+	3	1	$+e$	$+1$
Σ^0			0	0
Σ^-			$-e$	-1
N^{*++}	4	$\frac{3}{2}$	$2e$	$+\frac{3}{2}$
N^{*+}			$+e$	$+\frac{1}{2}$
N^0			0	$-\frac{1}{2}$
N^{*-}			$-e$	$-\frac{3}{2}$



charge on each shelf. For nucleons—the oldest acquaintances of the baryon family—this mean charge of the multiplet is equal to $+1/2$. We shall use this value to compare the mean charges of all the other baryon multiplets (to obtain integral instead of fractional values, it is convenient to take the doubled difference of the mean charges of the multiplets). To illustrate, take the Ξ-particles. The mean charge is $-1/2$. The doubled difference between this and the nucleonic mean charge is -2. For a Σ-triplet we get (in the same way) the value -1, as for the Λ^0-particle. It is simple to perform similar calculations for mesons, taking pi-mesons as the standard (that is, taking their displacement as zero).

Do these figures bring anything to mind? Why, of course, they are the values of strangeness that played such an important role in the story about the weak interaction.

Interesting, indeed! Approaching the problem from another angle, that of particle systematics (we are not dealing with disintegrations or other mutual transformations), we again find it necessary to introduce strangeness.

Building blocks within building blocks

Confronted by the problem of putting our expanding collection of particles in order, we have chosen four features: spin, mass, charge and strangeness. The spin deter-

mines into which cabinet—baryon or meson—the particle is to be put (we do not consider leptons), the other quantities indicate the number of the shelf and its pigeonhole.

Now that everything is in order with complete systematics, of what use is it? Does it have any profound physical meaning? Have the characteristics been chosen properly to substantiate the classification? Imagine that instead of classifying particles we were dealing with biological systematics and took the weight of the animal as a basis. It might be that our closest relative would be a crocodile or a pig. It is not a question of whether this flatters us as human beings. Simply, systematics of that kind does not deepen our understanding of the animal kingdom.

In short, do we have a good classification of the elementary particles? Note first of all that all the characteristics are quantities that do not change under strong interactions. No matter what transformations occur due to the strong interaction, the electric charge of the original products is the same as that of the final products. The very same may be said of strangeness and spin. (Mass presents a more complicated problem, but we shall not deal with it here.)

The impression one gets from the foregoing is that there must be some kind of selection of material bearers of charge, strangeness, and spin, some kind of subparticles, which merge in specific combinations to form baryons and mesons and which do not appear or disappear in any transformations, but simply pass from one combination into another. If that is the situation, then conservation of charge or strangeness is no more remarkable than conservation of the number of parts in a child's erector set, irrespective of whether the parts go to make a locomotive or a windmill.

Elementary particles have long been poetically termed the building blocks of matter. But if there are subparticles, then these building blocks must consist of some other still more elementary building units. The idea is rather tempting, to say the least.

Quarks?

Firstly, let it be noted that our subparticles must have spin 1/2. Indeed, halves can be used to build integral

or half-integral spins, which is something we would not be able to do if we had at our disposal only building material with spin zero, unity or any other whole number.

But what about charge, mass and strangeness?

Here we are in for a surprise. The most harmonious picture is obtained, it appears, if we give up our habit of ascribing integral charges (that is, multiples of the electronic charge) to our building blocks within blocks, and use fractional charges.

Fractional charges! A short time ago this idea would have seemed preposterous. Yet three charges of this kind were introduced by Gell-Mann. He called them quarks.

How many quarks do we need? The fewer the better, naturally. The necessary minimum turns out to be three. In the literature, they are denoted by the letters p, n and λ (do not confuse with the symbols for the proton, neutron and lambda particle, which henceforth are designated by P, N and Λ).

As we have already agreed to do, the spins of all quarks will be taken as equal to 1/2, while the other properties can conveniently be set out in a table.

Symbol of quark	Electric charge	Strangeness	Baryon charge
p	$+2/3$	0	1/3
n	$-1/3$	0	1/3
λ	$-1/3$	-1	1/3

Now let us try to combine the quarks so as to obtain known particles. We can begin with the proton P. The proton has strangeness zero; hence, we confine ourselves to a set made up of p and n. A proton should consist of a total of three quarks, so that the total baryon number should come out to unity. If we further take into account that the charge of P is $+1$, we arrive at the only possible set: ppn. Incidentally, it is not the only one, for we forgot about spin. It is important that the total spin should equal 1/2. This can be achieved if we take it that the spins of two p-quarks are parallel and the spin of n is antiparallel to them (or to the z-axis, as physicists say, which means to some chosen direction in space). This

can be symbolized as follows:

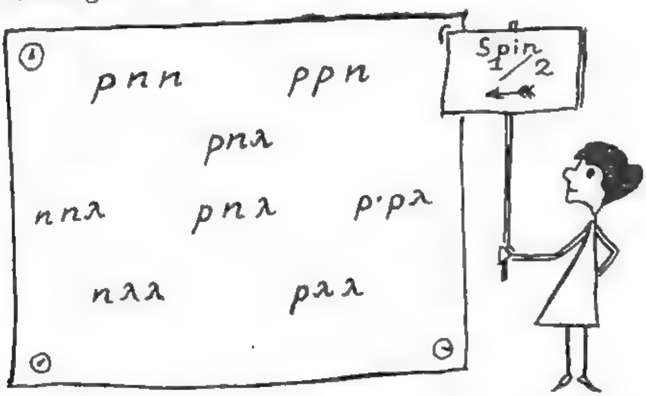
$$P = p\uparrow p\uparrow n\downarrow.$$

The arrows to the right of each symbol indicate the sense of its spin.

The recipe for forming particles out of quarks is now quite clear and does not appear to be complicated. Let us now try to arrange the quark triplets and see what we can get.

Let us first take combinations in which the spin is equal to 1/2. Which means that the spins of all quarks cannot have the same sense. A finer analysis based on the Pauli principle, which we discussed elsewhere in this book, indicates that not all spin combinations are allowed. For instance, if the total spin is 1/2, combinations of three identical quarks must be excluded. There are some other fine points that we shall not go into here. Let us simply write down the "permitted" combinations in rows. Inside each row, let the electric charge increase from left to right, and let the strangeness remain the same in each row; also let it diminish by unity when passing from one row to another one under it.

We then get the following table: *



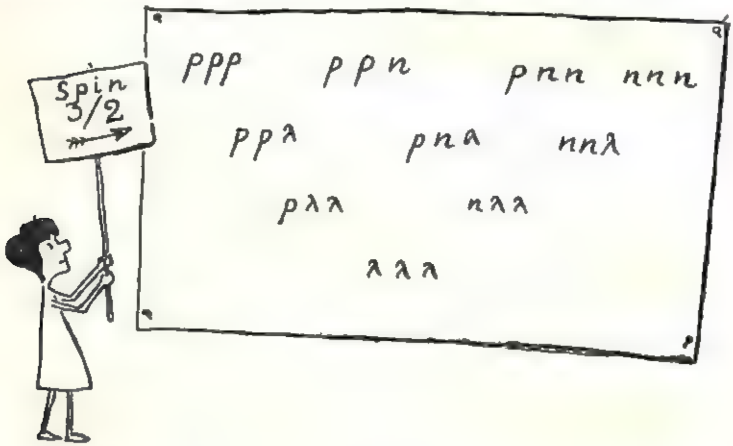
* The fact that we obtained two different combinations $pn\lambda$, was due to the possibility of different orientations of spin in this triplet. The Pauli principle does not allow two particles of the same kind to be in the same state at one time. But here all quarks are different, so the Pauli principle does not operate.

For the z-axis direction let us choose the direction of p. This will enable us to put an upward arrow for the first symbol of our triplet: $p\uparrow$. Now we must go through all combinations that yield total spin 1/2. There are obviously only two: $p\uparrow n\uparrow\lambda\downarrow$ and $p\uparrow n\downarrow\lambda\uparrow$. That explains why $pn\lambda$ appears twice in our table.

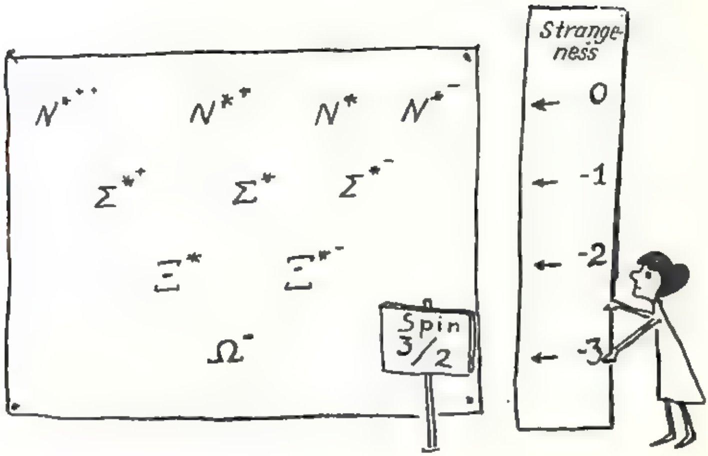
Now let us try to relate these triplets to particles. The combination *ppn* is the familiar proton. It is easy to see that *pnn* corresponds to the neutron. This means that the first row is a nucleon doublet. The charge singlet $pn\lambda$ may be correlated with the Λ^0-particle; the third row yields the triplet $\Sigma^- \Sigma^0 \Sigma^+$ and, finally, the last row contains combinations of quarks which correspond exactly to the Ξ-doublet, as regards strangeness and electric charges.

The result is definitely good. We have already constructed what is known as a baryon octet.

Let's continue this exciting building game and try quark triplets with total spin 3/2. Observing the same hules as in the foregoing case, let us build another table.



The first impression of this table is that we have constructed particles that are never found in nature. What is that *ppp* particle, for instance? Its electric charge should be +2, yet neither the proton, nor the Σ, nor the Ξ-particle has charge greater than unity. And the resonances? We have forgotten the resonances, yet they too must be brought into the general scheme of our classification. Among the resonances is the particle we need—the first of the resonance family: N^{*++}. Now we have a place for it in our systematics. Places will be found for the other resonances as well. Let's start out with the following table:



A predictea meeting

This is a rather familiar table. The asterisks show that we are dealing with resonances which are the excited states of particles with the corresponding symbols: the plus and minus signs indicate electric charge. But there is something new here too, the symbol Ω^-. A little over a year ago it was not mentioned in the most detailed surveys of elementary particles. Only a few enthusiastic theoreticians—those who believed in the new systematics—insisted on the existence of this particle. "Look for it," they said. They even gave a detailed description: charge -1, strangeness -3, baryon number $+1$, spin equal to 3/2. Even the mass was theoretically predicted. How? Very simply (everything is simple in afterthought).

A glance at the table of resonances will show us that the masses increase downwards from row to row. At the same time, one λ-quark is added as we go downwards from one row to the next. Lambda is absent in the top row, there is one λ in the next, two in the next one down, and, finally, Ω^- consists of three λ-quarks. If we correlate the growth in mass with the growth in number of quarks, we will find that *p*- and *n*-quarks (to which we ascribe identical masses) are lighter than the λ-quark. We can even estimate how much lighter. All we have to do is

compare resonance masses in neighbouring rows. The result will be a difference of about 0.16 nucleon mass (or 146 MeV in the presently accepted energy units). Thus, Ω^- should have a mass 146 MeV greater than the Ξ^*-particle. Returning to the table we find that the mass of Ξ^* is 1,530 MeV. Consequently, Ω^- must have a mass of 1,676 MeV. That is what theoreticians predicted.

On January 31, 1964, this particle was discovered experimentally! The Brookhaven laboratory in the United States was conducting a study of K^--particle collisions with protons. The Ω^--particle was detected in the reaction

$$K^- + P \rightarrow \Omega^- + K^+ + K^0.$$

In about 10^{-8} sec after it is born, the Ω^--particle decays as follows: $\Omega^- \rightarrow \Xi^0 + \pi^-$. The relatively long lifetime is due to the fact that decay according to strong patterns (that is, due to strong interactions) is prohibited by the conservation of strangeness; as we have already pointed out, change in strangeness is possible only in disintegrations due to the weak interaction.

The discovery of the Ω^--particle and the amazing precision with which theoretical predictions were corroborated could not but create a very strong impression. If systematics had earlier been regarded as an elegant, ingenious, but not very convincing playground for the imagination, feelings among physicists changed radically when the Ω^--particle was discovered. Big events were in the offing.

Advances galore

The new classification brought one success after another. Mesons and boson resonances fitted into the general system very naturally and easily. Bosons are particles with integral spin. Which means that they may be built up out of an even number of quarks. To be more precise, out of an even number of quarks and antiquarks*, for we want to get particles with a baryon charge of zero.

---

* Antiquarks are denoted by the quark symbols with a bar on top.

(In the case of antiquarks, the baryon charge, like the other quantum numbers, is opposite in sign to the charge of the corresponding quarks.)

The simplest combinaitons of this kind have the forms: $p\bar{p}$, $n\bar{n}$, $\lambda\bar{\lambda}$, $p\bar{n}$, $\bar{p}\lambda$, etc. If the spins are antiparallel, the resulting particles have zero spin. For example, the positive pion should be regarded as a combination $p\downarrow\bar{n}\downarrow$. True enough, the charge comes out to $2/3 + 1/3 = 1$, and the strangeness is zero, as it should be for a positive pion. If you want to build a strange meson, say K^+, take the combination $p\bar{\lambda}$. The electric charge of such a combination is unity, the strangeness, plus unity. To obtain bosons of integral spin, choose quark-antiquark pairs with spins in the same direction.

Don't bother writing it all out. We are sure the reader will believe us when we say that mesons and boson resonances fit into the framework of this systematics just as neatly as baryons.

And that is not all that the theory has achieved. Assuming that quarks have magnetic moments proportional to their charges, we can find a relationship between the magnetic moments of the particles made up of these quarks; for instance, we can find the ratio of the magnetic moments of the proton and neutron. It comes out to 3/2, while experiment yields 1.46. Brilliant agreement.

Difficulties arise

Just when the quarks were resolving problems and confidence in these new particles was growing, we would have to come up against some more difficulties. Such startling success, even triumph, such a brilliant prediction of a new particle, even if one prefers to disregard the systematization of known particles, and such an elegant explanation of the properties of these particles, such as the relationships between their magnetic moments—and then suddenly "difficulties".

What kind and where?

When Gell-Mann was asked whether quarks existed (their introduction into science was largely his doing), he said "Who knows?"

True, we don't know as yet. Which is much like the situation with the vector boson—the carrier of the weak interaction—about which we know almost everything,

except whether it actually exists or not. Perhaps the rational kernel that is so evident in particle systematics should not have been interpreted so quickly in the spirit of pictorial models of building blocks within blocks.

Quark adherents seek confirmation of their views even in historical parallels. For one thing, they point to the molecular theory that was formulated before direct proof was obtained of the existence of molecules. Then one could say that phenomena occur "as if molecules existed". But today we know that they do exist.

True, this parallel could be countered with other no less instructive cases. Maxwell constructed electrodynamics on the basis of a light-carrying ether, but we know today that there is no such ether. Fourier derived the presently used equations of heat conduction in the belief that heat was conveyed by some sort of ubiquitous fluid, yet caloric is never mentioned any more, except perhaps by historians.

Historical parallels are tricky things.

But why resort to such mebulous and unconvincing arguments instead of simply examining experiments? Couldn't we try to detect quarks experimentally? All the more so since these particles should, due to their fractional charge, be stable in the free state and prominent on the background of other particles. The search for quarks has begun. However, it would seem easier to create them in collisions brought about in particle accelerators. Simple indeed! If all particles consist of quarks, take any particle and bombard it with a suitable projectile and quark fragments should spatter in all directions.

The idea is simple, but the execution isn't. Our atom-smashing machines are not powderful enough. But nature has the accelerators we need. Once in a while cosmic-ray particles plunge into the earth's atmosphere with truly cosmic velocities. If such super-powerful particles collided with atomic nuclei present in the atmosphere, we should obtain some free quarks. We have already suggested that free quarks should be stable particles and after being born should behave something like this: quarks surrounded by "ordinary" particles attracted to them (these particles, incidentally, cannot completely neutralize them electrically since they have integral charges) form relatively large "blobs", which drop to the surface

of the earth or land in the ocean. This is important. If they hit a water surface, they should sink to the bottom and form fairly large concentrations of quarks. Then all one would have to do is scoop up this bottom water and evaporate it to get a sort of quark sludge.

No experiments of this kind have been performed. What is more, we are not sure that it is so easy to tell the quarks how to behave. Do we really know that much about these hypothetical particles? So far we have built quark models of familiar elementary particles more or less in the way a child builds with blocks. But the child's structure is stationary, while quarks are constantly on the move. The systems we are attempting to construct must be dynamic because—if for no other reason—the general principles of quantum mechanics prohibit particles closed in a small volume from being stationary. Quarks are in constant motion. We did not try to include their movements in our classification, but attempts are being made to do something in this respect and some advances have been registered, though the general situation with regard to quark motions is still very hazy.

Problems of the internal dynamics of quarks are closely tied in with yet another fundamental problem, that of the forces which maintain quark systems. That interquark interactions exist goes without saying. We even know that the interaction energy is very great. Otherwise, a relatively slight blow would knock any particle into its constituent quarks. The quark interaction energies must be colossal, hence also the mass defect. If in atomic nuclei the mass defect amounts to a fraction of one per cent of the total mass, then in particles made up of quarks it is close to one hundred per cent. The mass of quarks is presently considered to be from 7 to 10 times that of the nucleon mass. The proton consists of three quarks, and their total mass is of the order of twenty to thirty nucleon masses. How can it be that parts weighing 30 units form a whole that weighs one unit? We know now. When a whole forms, energy equivalent to 29 mass units is released. The binding energy is the same. This is a fantastic energy, far beyond anything we know at present.

But what forces cement quarks together that way? What is the nature of interquark interactions? And are

BARYON RESONANCES*

State	Mass, MeV	Resonance width, MeV	J^P	I, S	Decay types (and their relative weights)
P	938.256 ± 0.005		$^1/_2{}^+$		
n	939.550 ± 0.005				
$N^*_{1/2}$	~ 1480	~ 240	$^1/_2{}^+$		πN
$N^*_{1/2}$	1518 ± 10	$\left\{\begin{matrix}56\\125\end{matrix}\right.$?	$^3/_2{}^-$		$\left\{\begin{matrix}\pi N : & \pi\pi N : & \pi N^*_{3/2}\ (1335)\\ \sim 80\% : & ? & : ?\end{matrix}\right.$
$N^*_{1/2}$	1688	100	$^5/_2{}^+$	1/2.0	$\left\{\begin{matrix}\pi N : & \pi\pi N : K\Lambda : \pi N^*_{3/2}\ (1236)\\ \sim 80\% : & ? \quad : ? : ?\end{matrix}\right.$
$N^*_{1/2}$	2190	~ 200	?		$\left\{\begin{matrix}\pi N\\ \sim 30\%\end{matrix}\right.$
$N^*_{1/2}$	2645 ± 10	230	?		πN
$N^*_{1/2}$	2700	~ 100	?		$\eta N : \pi N$ $\sim 6\%$
$N^*_{3/2}$	1236 ± 2	125	$^3/_2{}^+$		πN
$N^*_{3/2}$	1924	170	$^7/_2{}^+$		$\pi N : \Sigma K : \pi N^*_{3/2}\ (1236)$
$N^*_{3/2}$	2360	~ 200	?	3/2.0	πN
$N^*_{3/2}$	2520		?		πN
$N^*_{3/2}$	2825	260	?		πN
$N^*_{5/2}$	1560 ± 20	220 ± 20	?	5/2.0	$p\pi^+\pi^+$
Λ	1115.40 ± 0.11		$^1/_2{}^+$		
Y^*_0	1405	50	?		$\left\{\begin{matrix}\Sigma\pi : \Lambda\pi\pi\\ 100 : (<1)\end{matrix}\right.$

Y_0^*	1518.9±1.5	16 ± 2	$^3/_2{}^-$	0, —1	Σπ : $\bar{K}N$: Λππ (55±7) : (29±4) : (16±2)
$\underline{Y_0^*}$	1660		$^1/_2$		Λη
Y_0^*	1815	70	$\underline{^5/_2\pm}$		$\bar{K}N$: Σπ : Λππ: Λη $\underline{80 : (<10) : (<15) : ?}$
Σ^+ Σ^0 Σ^-	1189.41±0.14 1192.4±0.3 1197.08±0.19		$^1/_2{}^+$		
Y_1^*	1382.1±0.9	53±2	$^3/_2{}^+$	1, —1	Λπ : Σπ (96±4) : (4±4)
Y_1^*	1660±10	44±5	$\underline{^3/_2{}^-}$		$\bar{K}N$: Σπ : Λπ : Λππ : Λππ 5 : 31 : 21 : 27 : 16
$\underline{Y_1^*}$	1765±10	60±10	$^5/_2{}^-$	1, —1	$\bar{K}N$: 60%
Ξ^0 Ξ^- Ξ^*	1314.3±1.0 1320.8±0.2 1529.1±1.0	7.5±1.7	$^1/_2{}^+$ $^5/_2{}^-$ $^3/_2\pm$	12, —2	Ξπ ~ 100%
$\underline{\Xi^*}$	1810±20	~ 70	?		$\Xi^*\pi$: Ξππ : Ξπ : $\Lambda\bar{K}$: $\Sigma\bar{K}$ 45 : (> 10) : 9 : 35 : 1.5 (< 50) : (< 50) : (10) 100 : ?
Ω^-	1675±3		?	0, —3	

* Question marks and the symbol ___ indicate "unreliable"; J is spin, P-parity; I-isotopic spin, and S-strangeness.

these interactions of one kind? Finally, in what relation do they stand to already known forces? For instance, do not nuclear forces, which we have interpreted as a manifestation of meson exchange by nucleons, reduce to some kind of special types of interaction between quarks, all the more so since nucleons consist of quarks and mesons (according to our model) are quark-antiquark pairs.

These and many other questions confront us. Put briefly, we must investigate the dynamical basis of the new particle systematics.

There is much that we still have to learn. For one thing, we cannot answer the question posed at the very start of the book: How many fundamental types of interaction are there, after all (a short time back we were sure there were only four).

We are positive that our grasp of nature has become more profound. These new ideas and the new classification of elementary particles, for one, (even if we disregard the remarkable discoveries described here) are of lasting value if for no other reason than they boldly pose specific questions and map out definite pathways. And even if tomorrow it is proved experimentally that there are no quarks at all in nature, this will not cross out the advances made in systematics, but will simply indicate that our models were only rough approximations to the great original—NATURE.

TO THE READER

Mir Publishers would be glad to have your opinion on the translation and the design of this book.

Please send all suggestions to Mir Publishers 2, Pervy Rizhsky Pereulok, Moscow, U.S.S.R.

Printed in the Union of the Soviet Socialist Republics

ABOUT THE AUTHORS

The authors of this book are Candidates of Physico-mathematical sciences Vladimir Grigoryev and Gennady Myakishev. In 1948 they graduated from the Physics Department of Moscow University and completed their graduate studies in 1951. At present both are Associate Professors of the Moscow University Physics Department.

V. Grigoryev has published over twenty papers on problems of quantum field theory and has been particularly interested in particle formation via high-energy collisions. G. Myakishev has written a number of works on electronics and problems of the methodology of science, several textbooks on elementary physics and numerous articles for popular-science magazines.

MIR PUBLISHERS · MOSCOW